

RESULT 1	AR399683	1161 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	AR399683				
DEFINITION	Sequence 1 from patent US 6620909.				
ACCESSION	AR399683				
VERSION	AR399683.1	GI:40142150			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
	Unclassified.				
REFERENCE	1 (bases 1 to 1161)				
AUTHORS	Piddington, C.S. and Bishop, P.D.				
TITLE	Antipocyte complement related protein homolog zacr2				
JOURNAL	Patent: US 6620909-A 1 16-SEP-2003;				
FEATURES	Location/Qualifiers				
	1..1161				
	/organism="Unknown"				
	/mol_type="genomic DNA"				
ORIGIN					

Db	241	AGCCTGCTGGCCCCCAGGGGCCACCGGGCCCCCAGAGAGCCCAAGGGCCTTCAGAAATG	300
QY	301	ATGGGAGCGAATGGGCTTCTCTGGCAAGACGGCCAAAGTAGCAACGACGGCGACCCGGGG	360
Db	301	ATGGGAGCGAATGGGCTTCTCTGGCAAGACGGCCAAAGTAGCAACGACGGCGACCCGGGG	360
QY	361	GACAGCGGAGAGGAAAGGTCCACTGGCCGGACAGGTAAACGGGGGAAAGCCAGACCAAG	420
Db	361	GACAGCGGAGAGGAAAGGTCCACTGGCCGGACAGGTAAACGGGGGAAAGCCAGACCAAG	420
QY	421	GGCAAAAGCCGGGGGCATTTGGGCGGGCTTGACCCCGTGGCCCCCAAGGGGGTCAACGTATC	480
Db	421	GGCAAAAGCCGGGGGCATTTGGGCGGGCTTGACCCCGTGGCCCCCAAGGGGGTCAACGTATC	480
QY	481	CCCGGAAAGCATGSCACACAGGCAAGAGGGGCTCAAAGGGCAAGAAAGGGAGCCAGGC	540
Db	481	CCCGGAAAGCATGSCACACAGGCAAGAGGGGCTCAAAGGGCAAGAAAGGGAGCCAGGC	540
QY	541	CTCCCAAGGCCCCCTGACAGTGTGGGCAATGCCATACAAAGTAGCTTTCTCGGTGGCAGTG	600
Db	541	CTCCCAAGGCCCCCTGACAGTGTGGGCAATGCCATACAAAGTAGCTTTCTCGGTGGCAGTG	600
QY	601	ACCAAGAGCTACCCACGGGAGCGGCTGCCATCAAGTTTGAACAAGTTCTGATGAAGAG	660
Db	601	ACCAAGAGCTACCCACGGGAGCGGCTGCCATCAAGTTTGAACAAGTTCTGATGAAGAG	660
QY	661	GGTGGCCACTCAAAATGCTTCCAGCGGCAAGTTGTTGTGCGGCGCTGGGATCTAATCAG	720
Db	661	GGTGGCCACTCAAAATGCTTCCAGCGGCAAGTTGTTGTGCGGCGCTGGGATCTAATCAG	720
QY	721	TTCACTTACGACATCAGCTGGGCCAACAGCACTGGGCATCGGCTGGTGTACAACAGGC	780
Db	721	TTCACTTACGACATCAGCTGGGCCAACAGCACTGGGCATCGGCTGGTGTACAACAGGC	780
QY	781	CAGTACCGCATCCGGACCTTTGATGTGCCAACACGGGCAACACAGATGTGTGACTCAGGCTCC	840
Db	781	CAGTACCGCATCCGGACCTTTGATGTGCCAACACGGGCAACACAGATGTGTGACTCAGGCTCC	840
QY	841	ACCATCTGTGGCTCTCAAGCAGGAGTGAAGAAATTTGGCTGCAGATCTTCTACTCAGACAG	900
Db	841	ACCATCTGTGGCTCTCAAGCAGGAGTGAAGAAATTTGGCTGCAGATCTTCTACTCAGACAG	900
QY	901	AACGGGCTTCTTATGACCTTACTGGAACAGACTCTTTACGGGCTTCTTAACTAT	960
Db	901	AACGGGCTTCTTATGACCTTACTGGAACAGACTCTTTACGGGCTTCTTAACTAT	960
QY	961	GGCGACCAAGATGACCCCAACGAGGTATGACATGSCACGGCGGTCTCCAGGAGGGAA	1020
Db	961	GGCGACCAAGATGACCCCAACGAGGTATGACATGSCACGGCGGTCTCCAGGAGGGAA	1020
QY	1021	CAAGCTTCTGACTTGGGCTTACAGAGCAAGACCCCAACTGTATGGCTGGGGGTGGGGG	1080
Db	1021	CAAGCTTCTGACTTGGGCTTACAGAGCAAGACCCCAACTGTATGGCTGGGGGTGGGGG	1080
QY	1081	GTCGAGTAGAGGGTTCATAGCTCAGAGCTCACTCCTCGCTCTTTTTCCTCCTCAAT	1140
Db	1081	GTCGAGTAGAGGGTTCATAGCTCAGAGCTCACTCCTCGCTCTTTTTCCTCCTCAAT	1140
QY	1141	AAATCCAAACCTTTTATTC	1161
Db	1141	AAATCCAAACCTTTTATTC	1161

RESULT 2	LOCUS	AX039965	AX039965	1161 bp	DNA	linear	PAT 18-NOV-2000
DEFINITION	Sequence 1 from Patent WO0063376.	AX039965					
ACCESSION	AX039965						
VERSION	AX039965.1						
KEYWORDS	GI:1122984						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Piddington,C.S. and Bishop,P.D.
TITLE	Adipocyte complement related protein homolog zacrp2
JOURNAL	Patent: WO 0063376-A 1 26-OCT-2000; ZymoGenetics, Inc. (us)
FEATURES	location/qualifiers 1..1161
SOURCE	/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 133..990 /note="unnamed protein product" /codon_start=1 /protein_id="CACI16634.1" /db_xref="GI:11229985" /translation="MIPPVLLAALPCADPLLGFARDPKSPOLVSLPGFQIGPGRAPGSGMMGRMGPPGDODHDGDEDSGESEPPPERTGRGRPKRGKAGAGRAPRPRKSNKTRPKHGCPKPKGKGPKEPRLPGPCSGSHHTSAFSVAVTKSPRESLPIFKDKILNMBEGHNINASSGKVLCVPGLYIFTYDITTLANKLATGLVHQQRIRFDPAANTGNHVASGSTJIALAKQDBEWMQLIFYESQNGLFPDYPTWLSLFTGPLIADQDDPNBV"
CDS	
ORIGIN	
Query Match	100.0%; Score 1161; DB 6; Length 1161;
Best Local Similarity	100.0%; Prid. No. 3.1e-213;
Matches 1161; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GGAAATCATATCCCTGGGGCGACGCCTGTGCCCGCTGTGCCTGAAGAAAAGCCGGAC 60
Dd	1 GGAAAACTATCTCTGGGGCCCAACGCTCTTGCCCCGCTGTGCCTGAGAAAAGCCGGAC 60
Oy	61 GCGAGCCCCCGCAGAGCTTCTTTGCTCGGAGCGCCCTGAGCGTGGCGGGACCGCG 120
Dd	61 GCGAGCCCCCGCAGAGACTTCTTTGCTCCGAGCGCCCTGAGCGTGGCGGGACCGCG 120
Oy	121 AGGGTAACCAACAATGATTCCTCTGGGTGCTCTTGCTGTGTGCCCTCCCCTGTGCTGTAC 180
Dd	121 AGGGTAACCAACAATGATTCCTCTGGGTGCTCTTGCTGTGTGCCCTCCCCTGTGCTGTAC 180
Oy	181 CCACGTGCTTGGCGCCTTGTCTCGAGGGGACTTCGCGAAAAGGCTCCCTCAAATGGTCTGC 240
Dd	181 CCACTGCTTGGCGCCTTGTCTCGAGGGGACTTCGCGAAAAGGCTCCCTCAAATGGTCTGC 240
Oy	241 AGCTGCTGTGCCCCCAGAGGCCCAACCAGGCCCCCAAGAGGCCCAAGGAGCCCTCAGAATG 300
Dd	241 AGCTGCTGTGCCCCCAGAGGCCCAACCAGGCCCCCAAGAGGCCCAAGGAGCCCTCAGAATG 300
Oy	301 ATGGGACGAATGGGCTTCTTGCGAAGAAGCGCCAAATGACACAGAGGCGACCGGGGG 360
Dd	301 ATGGGACGAATGGGCTTCTTGCGAAMAACGGCCAAATGACACAGAGGCGACCGGGGG 360
Oy	361 GACAGCGGAGGAAGGTCCAACCTGGCCGACAGGTAACTGGGGGAAAACCAAGCAAAAG 420
Dd	361 GACAGCGGAGGAAGGTCCAACCTGGCAAGGTAACTGGGGGAAAACCAAGCAAAAG 420
Oy	421 GGCAGAACCGGGGCAATTGGCGCGGCTTGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACC 480
Dd	421 GGCAGAACCGGGGCAATTGGCGCGGCTTGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACC 480
Oy	481 CCCGGGAAGCATGGCAACAAGGCAAGAAGGGGCCMAAGGGGCAAGAAAGGGGAGCCAGAGC 540
Dd	481 CCCGGGAAGCATGGCAACAAGGCAAGAAGGGGCCMAAGGGGCAAGAAAGGGGAGCCAGAGC 540
Oy	541 CTCCCAAGGCCCCCTGACGCTGTGGCAAGTGCATACCAAGTCAAGCTTTTCTCGGTGGCAATG 600
Dd	541 CTCCCAAGGCCCCCTGACGCTGTGGCAAGTGCATACCAAGTCAAGCTTTTCTCGGTGGCAATG 600
Oy	601 ACCAAGAGCTTACCAACGGGAGCGGCTGGCCATCAAGTTTGAACAAGATTCTGATGAACAG 660
Dd	601 ACCAAGAGCTTACCAACGGGAGCGGCTGGCCATCAAGTTTGAACAAGATTCTGATGAACAG 660
Oy	661 GGTGGCCACTCAATGCTTCCAGCGGCAAGTTGCTGCGGCGTGCCTGGGATCTACTAC 720

Db 661 GGTGGCCATCAATGCTTCCAGCGGCAAGTTGCTGCGGCTGCTGGATCTACTAC 720  
Qy 721 TTCACTTACGATCAGCTGGCCCAACAGCACTGGCCATGCGCTGGTGCACAGCGC 780  
Db 721 TTCACTTACGATCAGCTGGCCCAACAGCACTGGCCATGCGCTGGTGCACAGCGC 780  
Qy 781 CAGTACCGCATCCGGAACCTTTGATGCAACCGGCAACCGATGAGGCTCAGGCTCC 840  
Db 781 CAGTACCGCATCCGGAACCTTTGATGCAACCGGCAACCGATGAGGCTCAGGCTCC 840  
Qy 841 ACATCTGCTCTCAAGCAGAGGTGACGAGTTGGCTGAGATCTTCTACTCAGAGAG 900  
Db 841 ACATCTGCTCTCAAGCAGAGGTGACGAGTTGGCTGAGATCTTCTACTCAGAGAG 900  
Qy 901 AACGGGCTCTTCTATGACCTTACTGACAGACAGCTCTTTACGGGCTTCTATATCTAT 960  
Db 901 AACGGGCTCTTCTATGACCTTACTGACAGACAGCTCTTTACGGGCTTCTATATCTAT 960  
Qy 961 GCGACACAGATGACCCCAACAGAGTATAGCATGCCAGGCGGCTCTCCAGAGGAGAA 1020  
Db 961 GCGACACAGATGACCCCAACAGAGTATAGCATGCCAGGCGGCTCTCCAGAGGAGAA 1020  
Qy 1021 CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGCCCAACTGTAGGCTGGGAGTGGAGG 1080  
Db 1021 CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGCCCAACTGTAGGCTGGGAGTGGAGG 1080  
Qy 1081 GTGAGTGAAGGAGTGTAGAGCTCAAGGCTCACTCTCGGCTCTTTTTCCTTCATTT 1140  
Db 1081 GTGAGTGAAGGAGTGTAGAGCTCAAGGCTCACTCTCGGCTCTTTTTCCTTCATTT 1140  
Qy 1141 AATCCAACTTTTATCA 1161  
Db 1141 AATCCAACTTTTATCA 1161

RESULT 3  
AX698016 1161 bp DNA linear PAT 02-Apr-2003  
LOCUS AX698016 Sequence 1 from Patent WO03010197.  
DEFINITION AX698016  
ACCESSION AX698016  
VERSION AX698016.1 GI:29499054  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1  
AUTHORS Lucas, J., Scallia, A. and Dialynas, D.  
TITLE Gng-1 polynucleotides and polypeptides and uses thereof  
JOURNAL Patent: WO 03010197-A 1 06-FEB-2003;  
GENSET (FR)  
FEATURES  
source 1. 1161  
Location/Qualifiers  
1. 1161  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1. 1132  
133..990  
/note="unassigned protein product"  
/codon\_start=1  
/protein\_id="CAD87410.1"  
/db\_xref="GI:29499054"  
/translation="MIPWVLLAALPCADPDLGAFARDFRKSPLVCSLPGRPGP  
PGPGRGPGSGMGRMGFPGKDGQDGDGSDGSGSPRTGNRKGPGKAGAI  
GRAPRPGKVGNTGKGTGPKGKPKGKBPGLPBCSGSGHTSASFVAVTKSY  
PRRLPIKPKDKILNBSGHYASGKFCVPGIYFTYDITLANKILAIGLVNTGY  
RIKRPDANTGNHVAASSTILALQGBVWLIQIRYSRNGLPFYPIYTDLSLTFGLIY  
ADQDENBIV"  
991..1161

ORIGIN 3' UTR  
Query Match 100.0%; Score 1161; DB 6; Length 1161;

Best Local Similarity 100.0%; Pred. No. 3,1e-213;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGAATACTATGCTGGGCGGAGCGCTTGCCCGGCTGCTGCGTGAAGAAAGCGGAGC 60  
Db 1 GGAATACTATGCTGGGCGGAGCGCTTGCCCGGCTGCTGCGTGAAGAAAGCGGAGC 60  
Qy 61 GGGAGGCGGCGGAGAGCTTCTTGCTCGGAGAGCGCCCTGGAGAGTGGCGGAGCGGCG 120  
Db 61 GGGAGGCGGCGGAGAGCTTCTTGCTCGGAGAGCGCCCTGGAGAGTGGCGGAGCGGCG 120  
Qy 121 AGGTAACAACATGATCCCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
Db 121 AGGTAACAACATGATCCCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
Qy 181 CCACTGCTGGCGCTTGTGCTCGAGGAGCTTCCGAAAGGCTCCCTCAATGCTGCTGC 240  
Db 181 CCACTGCTGGCGCTTGTGCTCGAGGAGCTTCCGAAAGGCTCCCTCAATGCTGCTGC 240  
Qy 241 AGCTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
Db 241 AGCTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
Qy 301 ATGGAGCAATGAGCTTCTGCGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
Db 301 ATGGAGCAATGAGCTTCTGCGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
Qy 361 GACAGCGGAGAGAGGTCCACTGCGGAGACAGTAAACCGGAGAAAGCGAGCCAAAG 420  
Db 361 GACAGCGGAGAGAGGTCCACTGCGGAGACAGTAAACCGGAGAAAGCGAGCCAAAG 420  
Qy 421 GGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 480  
Db 421 GGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 480  
Qy 481 CCCGGAAGATGGCACAACAGGCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
Db 481 CCCGGAAGATGGCACAACAGGCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
Qy 541 CTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600  
Db 541 CTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600  
Qy 601 ACCAAGAGCTAACCAAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660  
Db 601 ACCAAGAGCTAACCAAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660  
Qy 661 GGTGGCACTTACATGCTTCCAGCGGCAAGTTCGTCGCGGCGGCGGCGGCGGCGGCGG 720  
Db 661 GGTGGCACTTACATGCTTCCAGCGGCAAGTTCGTCGCGGCGGCGGCGGCGGCGGCGG 720  
Qy 721 TTCACTTACGATCAGCTGGCCCAACAGCACTGGCCATGCGGCTGGTGCACAGCGC 780  
Db 721 TTCACTTACGATCAGCTGGCCCAACAGCACTGGCCATGCGGCTGGTGCACAGCGC 780  
Qy 781 CAGTACCGCATCCGGAACCTTTGATGCAACCGGCAACCGATGAGGCTCAGGCTCC 840  
Db 781 CAGTACCGCATCCGGAACCTTTGATGCAACCGGCAACCGATGAGGCTCAGGCTCC 840  
Qy 841 ACATCTGCTCTCAAGCAGAGGTGACGAGTTGGCTGAGATCTTCTACTCAGAGAG 900  
Db 841 ACATCTGCTCTCAAGCAGAGGTGACGAGTTGGCTGAGATCTTCTACTCAGAGAG 900  
Qy 901 AACGGGCTCTTCTATGACCTTACTGACAGACAGCTCTTTACGGGCTTCTATATCTAT 960  
Db 901 AACGGGCTCTTCTATGACCTTACTGACAGACAGCTCTTTACGGGCTTCTATATCTAT 960  
Qy 961 GCGACACAGATGACCCCAACAGAGTATAGCATGCCAGGCGGCTCTCCAGAGGAGAA 1020  
Db 961 GCGACACAGATGACCCCAACAGAGTATAGCATGCCAGGCGGCTCTCCAGAGGAGAA 1020  
Qy 1021 CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGCCCAACTGTAGGCTGGGAGTGGAGG 1080

Db 1021 CAAGCTTCTGGACTTGGGCTTACAGAGCAAGACCACAACCTAGGCTGGGGGTGGGGG 1080  
Qy 1081 GTGAGGAGAGCGGTTTACGCTCAGGCTCAGCTCTCGCTCTTTTTCCTCCCTTCAAT 1140  
Db 1081 GTCAAGAGAGCGGTTTACGCTCAGGCTCAGCTCTCGCTCTTTTTCCTCCCTTCAAT 1140  
Qy 1141 AAATCCAAACCTTTTATTC 1161  
Db 1141 AAATCCAAACCTTTTATTC 1161

RESULT 4  
AF329836 1176 bp mRNA linear PRI 12-MAR-2001  
LOCUS Homo sapiens complement-c1q tumor necrosis factor-related protein  
DEFINITION (C1RP2) mRNA, complete cds.  
ACCESSION AF329836  
VERSION AF329836.1 GI:13274517  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1176)  
AUTHORS Piddington,C.S. and Bishop,P.  
TITLE Homo sapiens complement-c1q tumor necrosis factor-related protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1176)  
AUTHORS Piddington,C.S. and Bishop,P.  
TITLE Direct Submission  
JOURNAL Submitted (20-DEC-2000) Bioinformatics, ZymoGenetics, Inc., 1201  
Baslake Ave. East, Seattle, WA 98102, USA

FEATURES  
source location/Qualifiers  
1..1176  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5q34"  
1..1176  
/gene="C1RP2"  
/note="ZACRP2"  
148..1005  
/gene="C1RP2"  
/codon\_start=1  
/product="complement-c1q tumor necrosis factor-related  
protein"  
/protein\_id="AAK17960.1"  
/db\_xref="GI:13274518"  
/translation="MIPVVLIAALPCADPLGAFARDFRKSPOLVGLPQPS  
PGPGAPGSGMMGRMGPQKIDQDHDGDRGDSBERGPRGRTGNRKA  
GRAGRPKPVNCTPGKGTGPKGKGGKGRPLPSPCCSGSHTSASVATKSY  
PRRLPTKFDKILNBSGHTNASKRVCGVPGIYFTYITILANKRLAIGLVHNGY  
RIKFDNANTGNHDVAGSTTILAKQDQVWMLQIFYSQNGLFYDPYWTDSLFTGFLY  
ADDDPENEV"

ORIGIN  
Query Match 100.0%; Score 1161; DB 9; Length 1176;  
Best Local Similarity 100.0%; Pred. No. 3,1e-213;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAAATATATGCTGGGGGCGAGCGCTCTGCGCGCTGCGCTGAGAAAAGCGGGAG 60  
Db 16 GAAAAATATATGCTGGGGGCGAGCGCTCTGCGCGCTGCGCTGAGAAAAGCGGGAG 75  
Qy 61 GCGGAGCGCGCGCGAGAGCTTTGCTCCGAGCGCCCTTGAGAGTGAGCGGCGAGCGCG 120  
Db 76 GCGGAGCGCGCGCGAGAGCTTTGCTCCGAGCGCCCTTGAGAGTGAGCGGCGAGCGCG 135  
Qy 121 AGGAGTACCAACATGATCCCTGGGTGCTCTGGCTGTGCTCCCTGTGCTGTGAC 180  
Db 136 AGGAGTACCAACATGATCCCTGGGTGCTCTGGCTGTGCTCCCTGTGCTGTGAC 195

Qy 181 CCAATGCTTGGCCCTTTTGCTCGAGGAACTTCCGAAAGGCTCCCTCAACTGATGTC 240  
Db 196 CCAATGCTTGGCCCTTTTGCTCGAGGAACTTCCGAAAGGCTCCCTCAACTGATGTC 255  
Qy 241 AGCTGCTTGGCCCTTCCGAGGCGCAACCGGCTCCCGAGAGCTCCAGGCTTCAAGATG 300  
Db 256 AGCTGCTTGGCCCTTCCGAGGCGCAACCGGCTCCCGAGAGCTCCAGGCTTCAAGATG 315  
Qy 301 ATGGAGCGAATGAGGCTTTCTGCAAGAGAGCGCAAGATGAGACGAGCGCGAG 360  
Db 316 ATGGAGCGAATGAGGCTTTCTGCAAGAGAGCGCAAGATGAGACGAGCGCGAG 375  
Qy 361 GACAGCGAAGAGAGAGTCCAGCTGCGCGGACAGGTAAACCGGGGAAAGCCAGAGCAAG 420  
Db 376 GACAGCGAAGAGAGAGTCCAGCTGCGCGGACAGGTAAACCGGGGAAAGCCAGAGCAAG 435  
Qy 421 GCGAAGCCCGGGCCATTGGGCGGGCTGCGCCCGTGGCCCAAGGGGGTCAACGTTACC 480  
Db 436 GCGAAGCCCGGGCCATTGGGCGGGCTGCGCCCGTGGCCCAAGGGGGTCAACGTTACC 495  
Qy 481 CCGGGAGAGATGGCAACAGGCGAAGAGGGGCCAAGGGGCAAGAGGGGAGCGAGG 540  
Db 496 CCGGGAGAGATGGCAACAGGCGAAGAGGGGCCAAGGGGCAAGAGGGGAGCGAGG 555  
Qy 541 CTCACAGGCGCTGACAGCTGTGGCAGTGGCCATACCAAGTCAAGCTTTCTGCGCAGTG 600  
Db 556 CTCACAGGCGCTGACAGCTGTGGCAGTGGCCATACCAAGTCAAGCTTTCTGCGCAGTG 615  
Qy 601 ACCAAGAGTACCAACGAGGAGCGGCTGCCATCAAGTTTGAACAAGTTCTGATGAACGAG 660  
Db 616 ACCAAGAGTACCAACGAGGAGCGGCTGCCATCAAGTTTGAACAAGTTCTGATGAACGAG 675  
Qy 661 GGTGGCCATCAATGTTTCCAGGGCAAGTTGCTGCGGCGCTGGGATCTACTAC 720  
Db 676 GGTGGCCATCAATGTTTCCAGGGCAAGTTGCTGCGGCGCTGGGATCTACTAC 735  
Qy 721 TTCACTACGACATCAAGCTGTGGCAACAAGCACTGCGCTGTGTGCAACAGG 780  
Db 736 TTCACTACGACATCAAGCTGTGGCAACAAGCACTGCGCTGTGTGCAACAGG 795  
Qy 781 CAGTACCGCATCCGAGCTTTGATGCCAACAACCGGCAACGATGTGCTCAGGCTCC 840  
Db 796 CAGTACCGCATCCGAGCTTTGATGCCAACAACCGGCAACGATGTGCTCAGGCTCC 855  
Qy 841 ACCATCTGGCTCAAGAGAGGGTGAAGAGTTTGGCTGAGATCTTCACTCAGAGCAG 900  
Db 856 ACCATCTGGCTCAAGAGAGGGTGAAGAGTTTGGCTGAGATCTTCACTCAGAGCAG 915  
Qy 901 AACGGCTCTTCTATGACCTTTACTGAGACAGAGCTTTTACGAGCTTCTATCTAT 960  
Db 916 AACGGCTCTTCTATGACCTTTACTGAGACAGAGCTTTTACGAGCTTCTATCTAT 975  
Qy 961 GCGGACGAGATGACCCCAAGAGATGACATGCGACGCGGCTCTCCAGGACGAGAA 1020  
Db 976 GCGGACGAGATGACCCCAAGAGATGACATGCGACGCGGCTCTCCAGGACGAGAA 1035  
Qy 1021 CAAGCTTCTGGACTTGGGCTTACAGAGCAAGACCCCAACTGAGGCTGGGGTGGGG 1080  
Db 1036 CAAGCTTCTGGACTTGGGCTTACAGAGCAAGACCCCAACTGAGGCTGGGGTGGGG 1095  
Qy 1081 GTGAGTGAAGCGGTTTCAAGCTCAGGCTCACTCTCCGCTCTTTTTCCTCCCTCAAT 1140  
Db 1096 GTGAGTGAAGCGGTTTCAAGCTCAGGCTCACTCTCCGCTCTTTTTCCTCCCTCAAT 1155  
Qy 1141 AAATCCAAACCTTTTATTC 1161  
Db 1156 AAATCCAAACCTTTTATTC 1176

RESULT 5  
AY358839 1167 bp mRNA linear PRI 03-OCT-2003  
LOCUS Homo sapiens clone DNA143501 C1RP2 (UN06349) mRNA, complete cds.  
DEFINITION



ACCESSION AY358839.1 GI:37182795  
VERSION AY358839.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1167)  
AUTHORS Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J., Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deneil,B., Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Haas,P.E., Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C., Schoenfeld,J., Seshagiri,S., Simons,L., Singh,J., Smith,V., Stinson,J., Vagstad,A., Vandlen,R., Watanabe,C., Weiland,D., Woods,K., Xie,M.H., Yamasu,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,W.I. and Godowski,P.  
The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: A Bioinformatics Assessment  
Genome Res. 13 (10), 2265-2270 (2003)  
JOURNAL  
PUBMED 12975309  
REFERENCE 2 (bases 1 to 1167)  
AUTHORS Clark,H.F.  
TITLE Direct Submission  
JOURNAL Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA  
FEATURES  
source  
1..1167  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DNA143501"  
1..1167  
/locus\_tag="UNQ6349"  
140..997  
/locus\_tag="UNQ6349"  
/notes="PRO21054"  
/product="CTRP2"  
/protein\_id="AA089198.1"  
/db\_xref="GI:37182796"  
/translation="MIPWVLLALPCADPLGAPARDRKSPOLVCSLPQSPGPGPRGAPSPGMRKMPGPGDGDHGDHGDGDSBSPRTNRKRPKRAKAI GRAPRPKGVNGTPGKHGTPGKXGPKGKGPGLPGPCSGSGSHSAVATKSY PRERLPKFDKILNBSGHVNASGKGVCGPIGYFTYDITLANKHAIQLVNIQGY RIRTFDANTGNHVDVAGSTIILAKQGDENVMLQIFYSQNGLFYDPYWTDSLFTFLIY ADQDDPNEV"  
ORIGIN  
Query Match 99.9%; Score 1160; DB 9; Length 1167;  
Best Local Similarity 100.0%; Pred. No. 4.8e-213;  
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGAATCATGAGCTGGGGCGGAGAGCTGCGGGCGGCTGCGGAGGAAAGCCGGAGC 60  
DB 8 GGAATCATGAGCTGGGGCGGAGAGCTGCGGGCGGCTGCGGAGGAAAGCCGGAGC 67  
QY 61 GCGAGAGCCCGCGAGAGAGCTTGTGCTCGGAGAGCCCTGAGAGTGGCGGGAGCCGCG 120  
DB 68 GCGAGAGCCCGCGAGAGAGCTTGTGCTCGGAGAGCCCTGAGAGTGGCGGGAGCCGCG 127  
QY 121 AGGATTAACACATGATCCCTGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAC 180  
DB 128 AGGATTAACACATGATCCCTGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAC 187  
QY 181 CCATGCTGGCGGCTTGTGCTGCGAGAGGAGCTTCCGGAAGGCTCCCTCAACTGCTGCTG 240  
DB 188 CCATGCTGGCGGCTTGTGCTGCGAGAGGAGCTTCCGGAAGGCTCCCTCAACTGCTGCTG 247  
QY 241 AGCTGCTGAGCCCGCGAGAGCCAGCCGCGCCCGCGAGAGCCAGAGGCGCTCTGAGATG 300  
DB 248 AGCTGCTGAGCCCGCGAGAGCCAGCCGCGCCCGCGAGAGCCAGAGGCGCTCTGAGATG 307

QY 301 ATGGAGCAATGGGCTTTCTCTGGCAAGAGCGGCCAAGATGGACACAGCGGAGCCGGGG 360  
DB 308 ATGGAGCAATGGGCTTTCTCTGGCAAGAGCGGCCAAGATGGACACAGCGGAGCCGGGG 367  
QY 361 GACAGCGGAGAGAGAGGTTCACCTGCGCGGACAGGTAAACGGGGAAACCAAGCAAG 420  
DB 368 GACAGCGGAGAGAGAGGTTCACCTGCGCGGACAGGTAAACGGGGAAACCAAGCAAG 427  
QY 421 GCGAAGCCCGGGGCAATTTGGCGGGGCTGGCCCCCGGTGGCCCAAGGGGGTCAACGTTAC 480  
DB 428 GCGAAGCCCGGGGCAATTTGGCGGGGCTGGCCCCCGGTGGCCCAAGGGGGTCAACGTTAC 487  
QY 481 CCGGAGAGAGATGGACACACAGGCAAGAAAGGGGCCCAAGGGGCAAGAAAGGGGCCAGG 540  
DB 488 CCGGAGAGAGATGGACACACAGGCAAGAAAGGGGCCCAAGGGGCAAGAAAGGGGCCAGG 547  
QY 541 CTCCAGAGCCCTGACAGTGGGAGTGGGACATACCAAGTCAAGCTTCTGCTGGTGGCAGTG 600  
DB 548 CTCCAGAGCCCTGACAGTGGGAGTGGGACATACCAAGTCAAGCTTCTGCTGGTGGCAGTG 607  
QY 601 ACCAAGAGCTTACCCACGAGCGGAGCGGCTGCCATCAAGTTTGAACAAGATTGTGATGAC 660  
DB 608 ACCAAGAGCTTACCCACGAGCGGAGCGGCTGCCATCAAGTTTGAACAAGATTGTGATGAC 667  
QY 661 GGTGGCCATTCATGATGCTTTCAGCGGCAAGTTGCTGCGGGCGGCTGGGGATCTACTAC 720  
DB 668 GGTGGCCATTCATGATGCTTTCAGCGGCAAGTTGCTGCGGGCGGCTGGGGATCTACTAC 727  
QY 721 TTCACTTACGACATCAAGCTGCGCAACCAAGCAAGTGGCGCTGCTGTCACCAAGCGC 780  
DB 728 TTCACTTACGACATCAAGCTGCGCAACCAAGCAAGTGGCGCTGCTGTCACCAAGCGC 787  
QY 781 CAGTACCGCATTCGAGACCTTTGATGCAACACCGGCAACCAAGTGGCGCTGAGGCTCC 840  
DB 788 CAGTACCGCATTCGAGACCTTTGATGCAACACCGGCAACCAAGTGGCGCTGAGGCTCC 847  
QY 841 ACCATCTGCTGCTCAAGAGAGGAGAGCAAGTTGGTGAAGATCTTCACTGAGAGCAG 900  
DB 848 ACCATCTGCTGCTCAAGAGAGGAGAGCAAGTTGGTGAAGATCTTCACTGAGAGCAG 907  
QY 901 AAGGGGCTTTTATGACCTTTACTGACAGAGAGCTCTTTTACGGGCTTCTATATCTAT 960  
DB 908 AAGGGGCTTTTATGACCTTTACTGACAGAGAGCTCTTTTACGGGCTTCTATATCTAT 967  
QY 961 GCCGACCAAGATGACCCCAAGAGATGACATGCAAGCGGCGGCTTCAAGCGAGGAA 1020  
DB 968 GCCGACCAAGATGACCCCAAGAGATGACATGCAAGCGGCGGCTTCAAGCGAGGAA 1027  
QY 1021 CAAGCTTCTGACTGGGCTTACAGAGACAGACCCCAACTGTAGGCTGGGGTGGGG 1080  
DB 1028 CAAGCTTCTGACTGGGCTTACAGAGACAGACCCCAACTGTAGGCTGGGGTGGGG 1087  
QY 1081 GTGAGTGAAGGAGTTCTAGGCTCAAGGCTCACTCTCGGCTCTTTTTCCTCTCAT 1140  
DB 1088 GTGAGTGAAGGAGTTCTAGGCTCAAGGCTCACTCTCGGCTCTTTTTCCTCTCAT 1147  
QY 1141 AATTCCAACCTTTTATTC 1160  
DB 1148 AATTCCAACCTTTTATTC 1167  
RESULT 6  
AX552234  
LOCUS AX552234 1171 bp DNA linear PAT 27-NOV-2002  
DEFINITION Sequence 143 from Patent WO0162927.  
ACCESSION AX552234  
VERSION AX552234.1 GI:25896472  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.



NIA-MGC Project URL: <http://mgc.nci.nih.gov>  
 On Aug 19, 2003 this sequence version replaced gi:15079785.  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Galtherburg, Maryland:  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@hgti.nih.gov](mailto:nisc_mgc@hgti.nih.gov)  
 Akhter, N., Ayale, C., Beckertov-Sternberg, S. M., Benjamin, B.,  
 Blakesley, R. W., Bouffard, G. G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N. L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
 Madhus, Q. L., Maiello, C., Masketi, B., Mastrian, S. D., McCloskey, J. C.,  
 McDowell, J., Pearson, R., Statistop, S., Thomas, P. J., Touchman, J. W.,  
 Young, A., Zhang, L.-H. and Green, E. D.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.linl.gov>  
Series: IRAL Plate: 27 Row: d Column: 17  
This clone was selected for full length sequencing because it passed the following selection criteria: Similarly but not identically to protein.

## Location/Qualifiers

```

source
1..2395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:19702 IMAGE:3530165"
/tissue_type="Muscle, rhadomyosarcoma"
/clone_lib="NIH MGC_17"
/lab_host="DH10B-R"
/vector="pOTB7"
1..2395
gene

```

## 121.

```

/gene="C1QTNF2"
/note="synonyms: CTRP2, zacrp2"
/db_xref="locustid:114898"
121_978
/gene="C1QTNF2"
/codon_start=1
/product="C1q and tumor necrosis factor related protein 2"
/protein_id="AAH1699.1"
/db_xref="GI:15079786"
/translat="MIRVILLCADPLILGAPARDPFKSPOLVCSLPDQSG
PGRPPAPEPSMGRMKPFPRKDDQDDHDDRDGDSGEGRPGPTGNRKGPKRGKAGAI
GREGKRGKGNGLPGKHGHPGKGGPKGKGEGLPGPSCSGHTKAFSLVAATKTS
PREHLPIFKDILINNEGHNASGKFCVGPPIYFYDITLARKHAIQILVINGQD
RIKTFDANTGNHDAVSGSTILALKQGDVWLQIPYSBQNLFPYDWTMSLFTSLIY
ADDDQPENV"

```

Query Match	99.0%	Score 1149;	DB 9;	Length 2395;
Best Local Similarity	100.0%	Pred. No. 5.8e-211;		
Matches 1149;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

133 ATGATCCCCCGGGTCTCCAGGACCTGTGCCCCCTGTGCTGACCCCACTGCTTGGC 192  
121 ATGATCCCCCGGGTCTCCAGGACCTGTGCCCCCTGTGCTGACCCCACTGCTTGGC 180

Oy		193	GCTTTTGTCGACAGGAATTCCGAAAGGCCTCCCTCAACTGGTCTGACGCTGTGCGC	252
Db		181	GCCTTTGTGCGAAGGACTTCGGAAAGGCTCCCCTCACTGGTCTGACGCTGTGCGC	240
Oy		253	CCCCAGGGGCCAACCCGGGCCCCCGAGAGGCCCAAGGCGCTCAAGAAATGATGGACAATG	312
Db		241	CCCCAGGGGCCAACCCGGGCCCCCGAGAGGCCCAAGGCGCTCTAGAATAATGGACGAATG	300
Oy		313	GGCTTTCTCTGGCAAAAGACGGCCAAAGATGACACGACGGCGCATCGGGGGGACAGCGAGAG	372
Db		301	GGCTTTCTCTGGCAAAAGACGGCCAAAGATGACACGACGGCGCATCGGGGGGACAGCGAGAG	360
Oy		373	GAAGTGTCACTTGGCCCGGACAGGTTAACCGGGGAAAAGCCAGGACCAAAGGGGCAAAGCCGGG	432
Db		361	GAAGTGTCACTTGGCCCGGACAGGTTAACCGGGGAAAAGCCAGGACCAAAGGGGCAAAGCCGGG	420
Oy		433	GCATTTGGGGCGGGCTGGGCCCCCGTAGGCCCCCAAGGGGGGTCAACGGTACCCCCGGAAAGCAT	492
Db		421	GCATTTGGGGCGGGCTGGGCCCCCGTAGGCCCCCAAGGGGGGTCAACGGTATCCCCGGGAAAGCAT	480
Oy		493	GGCACACCAAGGCANAGNAGGGGCCCAAGGGGCAAGAAAGGGAGCCAGGCTTCCAGGCCCC	552
Db		481	GGCACACCAAGGCANAGNAGGGGCCCAAGGGGCAAGAAAGGGAGCCAGGCTTCCAGGCCCC	540
Oy		553	TGCAGCTGTGGCAGTGGGCAATPACCAAGTCAGCTTCTTCGGTAGGACGTGACCAAGAGCTAC	612
Db		541	TGCAGCTGTGGCAGTGGGCAATPACCAAGTCAGCTTCTTCGGTAGGACGTGACCAAGAGCTAC	600
Oy		613	CCACGGGAGCGGGCTGCCCATCAAGTTTGAACAAGTTCTGATGAAAGAGGGTGGCCAATTAC	672
Db		601	CCACGGGAGCGGGCTGCCCATCAAGTTTGAACAAGTTCTGATGAAAGAGGGTGGCCAATTAC	660
Oy		673	AATGCTTCCAGCGGSCAAGTTGCTGTGGGGGGTGGCTGGGAATCTACTACTTCAACCTACGAC	732
Db		661	AATGCTTCCAGCGGSCAAGTTGCTGTGGGGGGTGGCTGGGAATCTACTACTTCAACCTACGAC	720
Oy		733	ATCACGGCTGGCCAAACAAGCACTTGGCCATCGGCTGGTGTGCAACAAGCCAGTAACCGCATC	792
Db		721	ATCACGGCTGGCCAAACAAGCACTTGGCCATCGGCTGGTGTGCAACAAGCCAGTAACCGCATC	780
Oy		793	CGGACCTTTTGAATGCAACAACCGGACACAAGATGTGGCTTCAGGCTCCACCATCTTGACT	852
Db		781	CGGACCTTTTGAATGCAACAACCGGACACAAGATGTGGCTTCAGGCTCCACCATCTTGACT	840
Oy		853	CTCAAGACGAGGTGACGAAGTTTGGCTGACATCTTCTACTACAGACGAAACGGGCTCTTC	912
Db		841	CTCAAGACGAGGTGACGAAGTTTGGCTGACATCTTCTACTACAGACGAAACGGGCTCTTC	900
Oy		913	TATNACCTTTATCTGGAACAGACAGGCTCTTTTACGGGCTTCCTAATCTATGCGGACCAAGAT	972
Db		901	TATNACCTTTATCTGGAACAGACAGGCTCTTTTACGGGCTTCCTAATCTATGCGGACCAAGAT	960
Oy		973	GACCCCAACGAGGTATAGACATGCAACGGCGGTCCTCCAGGACGAGGAACAAGCTTTGGA	1032
Db		961	GACCCCAACGAGGTATAGACATGCAACGGCGGTCCTCCAGGACGAGGAACAAGCTTTGGA	1020
Oy		1033	CTTGGGCTTACAGAGCAAGACCCCAACACTGTAGGCTGGGGGTGGGGGTGCGATGAGCG	1092
Db		1021	CTTGGGCTTACAGAGCAAGACCCCAACACTGTAGGCTGGGGGTGGGGGTGCGATGAGCG	1080
Oy		1093	GTTCCTAGCCTCAGGCTCACTCTCCGCTCTTTTTTTCCTTCATTAAATCCMAAAGCT	1152
Db		1081	GTTCCTAGCCTCAGGCTCACTCTCCGCTCTTTTTTTCCTTCATTAAATCCMAAAGCT	1140
Oy		1153	TTTTTATTC A 1161	
Db		1141	TTTTTATTC A 1149	

RESULT 8			
BC054506	1254 bp	mRNA	linear
LOCUS			
DEFINITION	BC054506	Homo sapiens Ctg and tumor necrosis factor related protein 2, mRNA	PRI 30-JUN-2004



Db 841 AGGTTGACGAGATTTGGCTGAGATCTTACTAGAGACAGACGGGCTCTTCTATAC 900  
Qy 920 CTTAAGGACAGACAGCTCTTTTACGGGCTTCTAATCTATGCCAGCAGATGACCCCA 979  
Db 901 CTTACTGACAGACAGCTCTTTTACGGGCTTCTAATCTATGCCAGCAGATGACCCCA 960  
Qy 980 ACGAGGTATAGACATGCGGCGGCTCTCCAGGCAAGGAACTTCTGACTTGGGC 1039  
Db 961 ACGAGGTATAGACATGCGGCGGCTCTCCAGGCAAGGAACTTCTGACTTGGGC 1020  
Qy 1040 TTACAGAGCAAGACCCCACTGAGGCTGAGGAGTGGGAGTGAAGAGCTTCTAG 1099  
Db 1021 TTACAGAGCAAGACCCCACTGAGGCTGAGGAGTGGGAGTGAAGAGCTTCTAG 1080  
Qy 1100 CTTGAGGCTGACCTCTGCGGCTCTTTTTCCTTCTTAAATCCAACTTTTATTT 1159  
Db 1081 CTTGAGGCTGACCTCTGCGGCTCTTTTTCCTTCTTAAATCCAACTTTTATTT 1140  
Qy 1160 CA 1161  
Db 1141 CA 1142

RESULT 9  
AC091842/c 152147 bp DNA linear PRI 23-OCT-2001

LOCUS AC091842 Homo sapiens chromosome 5 clone CTD-2155B8, complete sequence.  
AC091842  
AC091842.3 GI:16328266

VERSION HTG.  
KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 152147)  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

REFERENCE 2 (bases 1 to 152147)  
Unpublished

JOURNAL DOE Joint Genome Institute.

REFERENCE 3 (bases 1 to 152147)  
Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

REFERENCE 4 (bases 1 to 152147)  
Submitted (23-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

JOURNAL On Oct 23, 2001 this sequence version replaced gi:14579723.

COMMENT Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center

Quality: Phrap Quality >=40 99.7% of Sequence;  
Reclimated total Number of Errors is 0.3.

FEATURES  
SOURCE Location/Qualifiers  
1. 152147

ORGANISM /organism="Homo sapiens"  
/mol\_type="genomic DNA"

Db /db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTD-2155B8"

ORIGIN  
Query Match 67.5%; Score 783.8; DB 9; Length 152147;  
Best Local Similarity 99.7%; Pred. No. 6.1e-141;  
Matches 785; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 375 AGGTCACTGCGGCGGAGTAAACCGGGAAGCCAGACCAAGGCAAGCGCGGC 434  
Db 16142 AGGTCACTGCGGCGGAGTAAACCGGGAAGCCAGACCAAGGCAAGCGCGGC 16083

Qy 435 CATTGGGCGGCTGGCCCCCGTGGCCCAAGGGGGTCAAGGTAACCCCGGGAAGCATGG 494  
Db 16082 CATTGGGCGGCTGGCCCCCGTGGCCCAAGGGGGTCAAGGTAACCCCGGGAAGCATGG 16023  
Qy 495 CACACAGGCAAGAGGGGGCCCAAGGCAAGAAAGGAGCAGGCTCCAGGCCCCCTG 554  
Db 16022 CACACAGGCAAGAGGGGGCCCAAGGCAAGAAAGGAGCAGGCTCCAGGCCCCCTG 15963  
Qy 555 CAGCTGGGCAAGTGGCCATACCAAGTCAAGTCTTCTCGGTGGCAATACCAAGTACCC 614  
Db 15962 CAGCTGGGCAAGTGGCCATACCAAGTCAAGTCTTCTCGGTGGCAATACCAAGTACCC 15903  
Qy 615 ACGGAGCGGCTGGCCATCAAGTTTGAAGAATTTGAAGAAGAGGGTGGCACTCA 674  
Db 15902 ACGGAGCGGCTGGCCATCAAGTTTGAAGAATTTGAAGAAGAGGGTGGCACTCA 15843  
Qy 675 TGCTTCAGGCGGCAAGTGGCTGGCGGCTGGCGGCTGGAGTCTACTACTTCACTACGAT 734  
Db 15842 TGCTTCAGGCGGCAAGTGGCTGGCGGCTGGCGGCTGGAGTCTACTACTTCACTACGAT 15783  
Qy 735 CACGCTGGCCAAAGAGCACTGGCATGGGCTGGTGCACACGGCCAGTACCGCATCG 794  
Db 15782 CACGCTGGCCAAAGAGCACTGGCATGGGCTGGTGCACACGGCCAGTACCGCATCG 15723  
Qy 795 GACCTTGATGGCCAAAGCGGCAAGCGGCAAGTGGGCTGGGCTGCAATCTGGGCT 854  
Db 15722 GACCTTGATGGCCAAAGCGGCAAGCGGCAAGTGGGCTGGGCTGCAATCTGGGCT 15663  
Qy 855 CAAGCAGGCTGACCAAGTTGGCTGGAGTCTTCTACTGAGAGCAAGCGGCTCTTCTA 914  
Db 15662 CAAGCAGGCTGACCAAGTTGGCTGGAGTCTTCTACTGAGAGCAAGCGGCTCTTCTA 15603  
Qy 915 TGACCTTACTGACAGACAGCCTCTTTACGGGCTTCTAATCTATGCCACAGATGA 974  
Db 15602 TGACCTTACTGACAGACAGCCTCTTTACGGGCTTCTAATCTATGCCACAGATGA 15543  
Qy 975 CCCCAAGGTTATGACATGGCAAGCGGCTGGGCTGGAGGAGCAAGGCTTGGACT 1034  
Db 15542 CCCCAAGGTTATGACATGGCAAGCGGCTGGGCTGGAGGAGCAAGGCTTGGACT 15483  
Qy 1035 TGGGCTTACAGAGCAAGCCCAACTGATGAGCTGGGCTGGAGTACGCGGT 1094  
Db 15482 TGGGCTTACAGAGCAAGCCCAACTGATGAGCTGGGCTGGAGTACGCGGT 15423  
Qy 1095 TCTAGCTCAGGCTCAGCTCTGCGGCTTTTTCCTTCAATTAATCAACCTTT 1154  
Db 15422 TCTAGCTCAGGCTCAGCTCTGCGGCTTTTTCCTTCAATTAATCAACCTTT 15363  
Qy 1155 TTATTTCA 1161  
Db 15362 TTATTTCA 15356

RESULT 10  
AC112191

LOCUS AC112191 153248 bp DNA linear PRI 23-APR-2002  
DEFINITION Homo sapiens chromosome 5 clone RP11-34P1, complete sequence.  
AC112191  
AC112191.2 GI:20270112

VERSION HTG.  
KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 153248)  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

REFERENCE 2 (bases 1 to 153248)  
Unpublished

JOURNAL DOE Joint Genome Institute.

REFERENCE 3 (bases 1 to 153248)  
Submitted (20-FEB-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCES	AUTHORS	TITLE	JOURNAL	COMMENT
3 (bases 1 to 153248)	DOE Joint Genome Institute and Stanford Human Genome Center.	Direct Submission	Submitted (23-APR-2002)	DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
			On Apr 23, 2002 this sequence version replaced gi:18767463.	

FEATURES	Location/Qualifiers
source	1. .153248

## ORIGIN

Query Match		67.5%;	Score 783.8;	DB 9;	Length 153248;	
Best Local Similarity		99.7%;	Pred. No. 6,1e-141;			
Matches	785;	Conservative	0;	Mismatches	2;	
			Indels	0;	Gaps	0;
QY	375	AGGTCACCTTGAGCGGACAGGTAACCGGGGAAAGCCAGAGACCAAGGGCAAAAGCCGGGCG	434			
Db	20117	AGGTCACCTTGAGCGGACAGGTAACCGGGGAAAGCCAGAGACCAAGGGCAAAAGCCGGGCG	20117			
QY	435	CATTGGGGGGGCTGGCCCCCGTGGCCCCCAAGGGGTCAACGGTACCCCGGGAAAGCATGG	494			
Db	201177	CATTGGGGGGGCTGGCCCCCGTGGCCCCCAAGGGGTCAACGGTACCCCGGGAAAGCATGG	20233			
QY	495	CACACCAAGGCAAGAAAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGCGTCCAGAGCCCGTGG	554			
Db	20237	CACACCAAGGCAAGAAAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGCGTCCAGAGCCCGTGG	20297			
QY	555	CAGCTGTGGCAGTGGCCATACCAAGTCAGCTTTCGTGGTGGCAGTGAACCAAGAGCTACCC	614			
Db	20297	CAGCTGTGGCAGTGGCCATACCAAGTCAGCTTTCGTGGTGGCAGTGAACCAAGAGCTACCC	20359			
QY	615	ACGGAGCGGCTGCCCATCAAGTTTGAACAAGTTCTGATGAACGAGGGTGGCCACTTACAA	674			
Db	20357	ACGGAGCGGCTGCCCATCAAGTTTGAACAAGTTCTGATGAACGAGGGTGGCCACTTACAA	20411			
QY	675	TGCTTCCAGGGGCAAGTGGTGGTGGGGGTGGCGTGGGATCTAATCTTCACTTACGACAT	734			
Db	20417	TGCTTCCAGGGGCAAGTGGTGGTGGGGGTGGCGTGGGATCTAATCTTCACTTACGACAT	20477			
QY	735	CACGTGGCCAAACAAGCACTTGGGCATGGGCTGGTGCACAACGGCCAGTACCGGATCCG	794			
Db	20477	CACGTGGCCAAACAAGCACTTGGGCATGGGCTGGTGCACAACGGCCAGTACCGGATCCG	20533			
QY	795	GACCTTTGATCCCAACACCGGCAACCAAGATGTGGCTTCAGGCTTCCACCATCTTGGCTT	854			
Db	20537	GACCTTTGATCCCAACACCGGCAACCAAGATGTGGCTTCAGGCTTCCACCATCTTGGCTT	20597			
QY	855	CAACCAAGGATACCAAGATTGGCTTGCAGATCTTCTACTCAAGACGAAACGGGCTTCTTA	914			
Db	20597	CAACCAAGGATACCAAGATTGGCTTGCAGATCTTCTACTCAAGACGAAACGGGCTTCTTA	20657			
QY	915	TGACCCCTTACTGACAGACAGCTCTTTACGGGGCTTCTAATCTATGCGACCAAGATGA	974			
Db	20657	TGACCCCTTACTGACAGACAGCTCTTTACGGGGCTTCTAATCTATGCGACCAAGATGA	20711			
QY	975	CCCCACGAGGTATGACATGCCACCGCGGTCTTCCAGGACGGGAAACAAGCTTCTGAACT	1034			
Db	20717	CCCCACGAGGTATGACATGCCACCGCGGTCTTCCAGGACGGGAAACAAGCTTCTGAACT	20777			
QY	1035	TGGGCTTACAGAGCAAGACCCCAACACTTGAAGCTGGGGGTGGGGGATCGAGTGAACGGT	1094			

Dd	20777	TGGCTTACAGAACCAAGCCCCAACACTGAGGTGGGGGTGGGGGGTGAGTAGAGCGST	20836
Oy	1095	TCTAGCCTCAGGGTCACCTCTCCGACTCTTTTTTTCCCCTCATTAATCCAAACCTTT	1154
Dd	20837	TCTAGCCTCAGGGTCACCTCTCTGCTCTTTTTTTCCCCTCATTAATCCAAACCTTT	20896
Oy	1155	TTATTCA 1161	
Dd	20897	TTATTCA 20903	

RESULT 11	
LOCUS	BC030324
DEFINITION	BC030324 1236 bp mRNA linear ROD 24-FEB-2004
ACCESSION	Mus musculus Clq and tumor necrosis factor related protein 2, mRNA
VERSION	(cDNA clone MGC:40933 IMAGE:575090), complete cds.
KEYWORDS	BC030324 BC030324 MGC:40933 MGC:40933
SOURCE	MGC.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

1 (bases 1 to 1236)  
Strasberg, R.L., Feilgenfeldt, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D.,  
Altshuler, S.F., Zeeberg, B., Butelov, K.H., Schaefer, C.F., Bat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Ditachenko, L., Marinina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stableton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,  
Schaefer, T.E., Brownstein, M.U., Uedl, T.B., Toshiyuki, S.,  
Carrinck, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gamarathne, P.H., Richards, S.,  
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulys, S.W.,  
Vallalon, D.K., Muzny, D.M., Sodergren, E., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, B., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherchenko, Y.,  
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, B.D.,  
Dickson, M.C., Rodigues, A.C., Grimmood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,  
Scherer, A., Schen, J.B., Jones, S.U. and Harris, M.A., Mammalian  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1236)  
Strasberg, R.  
Submitted (07-MAY-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Contact: [nisc.mgc@hghl.nih.gov](mailto:nisc.mgc@hghl.nih.gov)  
Web site: <http://www.nisc.nih.gov/>

Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Khong, P., Latic, P., Legaspi, R.,  
Maduro, Q.L., Masillo, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W.,  
Tsungueon, C., Vogt, J.L., Walker, M.A., Wettersby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.







Db 582 CTCACAGGCCCCCTGAGCTGCGGAGTAGCCGAGCAATGCGCTTTTCGATGCGGTAA 641  
Qy 601 ACCAAGAGCTACCCACGCGGAGCGCTGCCATCAAGTTTGACAGATTTCTGATGAACAG 660  
Db 642 ACCAAGAGTTACCCACGAGCGAGCTGCCCATCAAGTTTGACAGATTTCTGATGAATAG 701  
Qy 661 GGTGGCCACTCAATAGCTTCCAGCGGCGCAAGTTGCTGCGGCGTGGCTGGGATCTACTAC 720  
Db 702 GGAAGCCACTCAATGATGATGATGAGGAGTTGCTGCTGAGCGTGGCCAGGAGTTCTATTAC 761  
Qy 721 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATG 780  
Db 762 TTATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 821  
Qy 781 CAGTACCGCATCCGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
Db 822 CAGTACCGCATCCGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 881  
Qy 841 ACATCTGCTGCTCAAGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
Db 882 ACATCTGCTGCTCAAGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 941  
Qy 901 AACGGGCTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db 942 AATGAGCTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001  
Qy 961 GCCACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1018  
Db 1002 GCTGATCAAG 1061  
Qy 1019 AACAGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1063  
Db 1062 ACTAGATTCGCGAGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1106

RESULT 14  
BD209734 1052 bp DNA linear PAT 17-JUL-2003  
LOCUS BD209734  
DEFINITION Compositions isolated from skin cells and methods for their use.  
ACCESSION BD209734  
VERSION BD209734.1 GI:33019504  
KEYWORDS JP 2002512798-A/206.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1052)  
Strachan,L., Sleeman,M., Watson,J.D., Onrust,R., Kumble,A. and  
Murison,J.G.  
Compositions isolated from skin cells and methods for their use  
Patent: JP 2002512798-A 206 08-MAY-2002;  
GENESIS RESEARCH AND DEVELOPMENT CORP LTD  
OS Rattus sp. (rat)  
PN JP 2002512798-A/206  
PD 08-MAY-2002  
PF 29-APR-1999 JP 2000546009  
PR 29-APR-1998 US 09/069726, 09-NOV-1998 US 09/188930 PT  
LORNA STRACHAN, MATTHEW SLEEMAN, JAMES DOUGLAS WATSON, RENE PI  
ONRUST,  
PI ANAND KUMBLE, JAMES GREG MURISON  
PC C12N15/09, A61K38/00, A61P9/00, A61P17/00, A61P29/00, A61P31/18, PC  
A61P35/00,  
PC C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12N15/ PC  
00,  
PC A61K37/02, C12N5/00  
CC Compositions isolated from skin cells and methods for their  
CC use.  
FT Key Location/Qualifiers  
FT source 1..1052  
FT /organism='Rattus sp. (rat)'.  
FT Location/Qualifiers

source 1..1052  
/organism='Rattus sp.'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:10118'

Query Match 60.4%; Score 701.6; DB 6; Length 1052;  
Best Local Similarity 82.9%; Pred. No. 7.4e-125;  
Matches 837; Conservative 0; Mismatches 169; Indels 4; Gaps 3;

ORIGIN

Qy 121 AGGGTAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
Db 34 AAGGTGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 93  
Qy 181 CCATGCTTGGCGCTTTGCTGCGAGGAGCTTCCGAAAGGCTCCCTCAATGATGATGATGATG 240  
Db 94 CCAATGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 153  
Qy 241 AGCGGCTGCGCCCGAGGCGCCACCGGCGCCCGAGGCGCCCGAGGCGCCCGAGGCGCCCGAGG 300  
Db 154 AGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 212  
Qy 301 ATGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 213 GTGGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 272  
Qy 361 GACAGCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
Db 273 GACAGTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 332  
Qy 421 GGCAGAGCGGAGGCTTGGGCGGAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480  
Db 333 GGCAGAGCTGAGGCTTGGGAGGCGGCTTGGGAGGCGGCTTGGGAGGCGGCTTGGGAGGCGG 392  
Qy 481 CCCGGAAGCATGAGCAACAGGAGAGAGGCGCCCAAGGCGCAAGAAAGGAGGAGCGAGC 540  
Db 393 CCCGGAAGCATGAGTATACGGGCAAGAGAGGAGCTTAAGGCGCAAGAAAGGAGGAGCGAGC 452  
Qy 541 CTCGAGGCGCTTGGAGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 600  
Db 453 CTCGAGGCGCTTGGAGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 512  
Qy 601 ACCAAGAGCTACCCAGGAGGAGGCGCTGCGCATCAAGTTTGACAGATTTCTGATGATGATG 660  
Db 513 ACCAAGAGTTACCACTGAGAGAGCTGCCATCAAGTTTGACAGATTTCTGATGATGATGATG 572  
Qy 661 GGTGGCCACTCAATAGCTTCCAGCGGAGAGTTGCTGCGGCGTGGCTGGGATCTACTAC 720  
Db 573 GGAAGCCACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 632  
Qy 721 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 780  
Db 721 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 740  
Qy 740 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 800  
Db 740 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 760  
Qy 760 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 820  
Db 760 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 780  
Qy 780 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 840  
Db 780 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 800  
Qy 800 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 860  
Db 800 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 820  
Qy 820 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 880  
Db 820 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 840  
Qy 840 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 900  
Db 840 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 860  
Qy 860 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 920  
Db 860 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 880  
Qy 880 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 940  
Db 880 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 900  
Qy 900 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 960  
Db 900 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 920  
Qy 920 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 980  
Db 920 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 940  
Qy 940 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 1000  
Db 940 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 960  
Qy 960 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 1020  
Db 960 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 980  
Qy 980 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 1040  
Db 980 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 1000  
Qy 1000 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 1060  
Db 1000 TTCACTTACGATCATCGCTGCGCAACAGCACTGCGCATCGGCTGATGATGATGATGATGATG 1020  
Qy 1020 AACAGCTTGGAGCTTGGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1079  
Db 933 CTAGAGATTCGCGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 992

QY 1080 GGTGAGTGAAGGTTCTAGCTCAGGCTCAGCTCTCCGCTCTTTT 1129  
Db 993 GTTCTTGGAGC--TTTATTCCAGGCAAGCTCTCTGTGTGCTTT 1040

RESULT 15  
AR341540 1052 bp DNA linear PAT 17-AUG-2003  
LOCUS Sequence 358 from patent US 6573095.  
DEFINITION AR341540  
ACCESSION AR341540  
VERSION AR341540.1 GI:33733675  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1052)  
AUTHORS Strachan, L.  
TITL Polynucleotides isolated from skin cells  
JOURNAL Patent: US 6573095-A 358 03-JUN-2003;  
FEATURES  
source 1. 1052  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 60.4%; Score 701.6; DB 6; Length 1052;  
Best Local Similarity 82.9%; Pred. No. 7.4e-125;  
Matches 837; Conservative 0; Mismatches 169; Indels 4; Gaps 3;

QY 121 AGGGTAACAACATGATCCCTGGAGTGTCTGTGCTGTGCTCCCTGTGTGCTGAC 180  
Db 34 AAGGTGACACACATGATCTCTGATGTCTTGGCTGTGCTTCCGTGTGTGCTGAC 93

QY 181 CCACTGCTTGGCGCTTTTGTCTGCAAGGACTTCCGAAAGGCTCCCTCACTGTCTGC 240  
Db 94 CCAATGCTTGTGTCTTGTCTGTGCTGCAAGGACTTCCGAAAGGCTTCACTGTGTGC 153

QY 241 AGCTGCTGTGCGCCCGAGGCGCAAGCGGCGCCCGAGGCGCCCGAGGCGCTCAGGAATG 300  
Db 154 AGTGTGCTGTGCTCCCAAGG--CAACTGTGCGCTTCAAGGACCAAGATCTTCAAGAAATG 212

QY 301 ATGGAGCAATGGGCTTCTGTGCAAAAGCGGCCAAGATGACAGACGGCGACCGGAGG 360  
Db 213 GTGGGAAGAATGGGTTTCTGTAGAGATGGCCAAAGCGGCGAGACCGAGACCGAGGG 272

QY 361 GACAGCGGAGAGAAAGTCCACTGTGCGGACAGGTAAACGGGGAAGCCAGACCAAG 420  
Db 273 GACAGTGAAGAAAGGTCCACTGTGCAAGGCAAGCAACGAGGAAACAAAGCAAG 332

QY 421 GGCAAAAGCGGAGCGCTATTTGGGCGGAGTGGCCCGGTGGCCCAAGGGGGTCAACGTACC 480  
Db 333 GGCAAAAGCTGGGCGCTATTTGGGAGCGGAGTCTTGAAGGACCAAGGGGGTCAAGTGTACC 392

QY 481 CCCGGGAAGCATGGCACACAGGCAAGAAAGGGGCCCAAGGCAAAAGGGAGCCAGGC 540  
Db 393 CCCGGGAACATGGTATATCCGGGCAAGAAAGGACCTTAAGGCAAAAGGGGAACCTGGG 452

QY 541 CTCCCAAGGCGCTGTGAGTGTGGCAAGTGTGCAATCAAGTCAAGTCTTCTGGTGGCAATG 600  
Db 453 CTCCCAAGCGCTGTGAGTGTGGCAAGTGTGCAAGTCAAGTCTTCTGGTGGCGTGA 512

QY 601 ACCAAGAGCTAACCAAGGAGCGGCTGCGCATCAAGTTTGAAGATTTCTATGACGAG 660  
Db 513 ACCAAGAGTTTACCAAGTGAAGCGACTGCGCATCAAGTTTGAAGATTTCTATGATGAG 572

QY 661 GATGGCCTTACCAATGTCTTCCAGCGGCAAGTTGTGTGCGGCGTGTGGATCTACTAC 720  
Db 573 GAGGGCCTTACCAATGTATTCAGTGGCAAGTTGTGTGCAAGTGTGCAAGGATCTATTAAC 632

QY 721 TTCACTTACGATCATGCTGTGGCAAGCAAGCACTGGCATGGCTGTGTGCAACAGGC 780  
Db 633 TTATCTATGATGATTAAGCTGTGGCAAGCAAGCACTGGCATGGCTGTGTGCAACATGGC 692

QY 781 CAGTACCGCATCCGAGCTTTTGAATGCCAACCGGCAACCAAGATGTGGCTCAGGCTCC 840  
Db 693 CAGTACCGCATCCGAGCTTTTGAAGCGCAACCGGCAACCAAGATGTGGCTCAGGCTCC 752

QY 841 ACCATCTGTGCTTCAAGCAAGGTTGACGAAGTTTGGCTGACATCTTCTACTAGAGCAG 900  
Db 753 ACCATCTGTGCTTCAAGCAAGGTTGATGATGATGATGATGATTTCTACTCGAGCAG 812

QY 901 AAGGGCTTCTTATGACCTTACTGAGCAGACAGGCTTTTACGGCTTCTTATCTAT 960  
Db 813 AATGAGCTTCTTATGACCTTATTTGACGAGCCTGTTCACCGGCTTCTATCTAC 872

QY 961 GCCACGAGATGACCCCAAGGATATGACATGACCAAGG--CGATCTTCCAGGCAAGGGA 1019  
Db 873 GCTGATCAAGAGAGCCCAATGAGGTATGACAAAGCTGGGGTTGAGGCTCAGGCAAGGGA 932

QY 1020 ACAAGCTTGTGACTTGGCTTACAGAGCAAGACCCCAACATGTAGGCTGGGGGTGGG 1079  
Db 933 CTAAGATTTCCGCAAGGCTGTGATGAGAAAGGATCTCTGAAGTGAAGGCTGGGGAAGTGGCA 992

QY 1080 GGTGAGTGAAGGTTTCTAGCTCAGGCTCAGCTCTCCGCTCTTTT 1129  
Db 993 GTTCTTGGAGC--TTTATTCCAGGCAAGCTCTCTGTGTGCTTT 1040

Search completed: December 25, 2004, 05:29:22  
Job time : 5282 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2004, 02:28:47 ; Search time 655 Seconds

(without alignments)  
9304.699 Million cell updates/sec

Title: US-10-621-787-1

Perfect score: 1161  
Sequence: 1 ggaataactagcctcgggcc.....aatccaactttattca 1161

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1161	100.0	1161	3 AAC65132	AAC65132 Human adi
2	1161	100.0	1161	8 AB224307	AB224307 Human GMG
3	1161	100.0	1161	11 ADM66421	ADM66421 CDNA enco
4	1161	100.0	1161	12 ADL26856	ADL26856 Human zac
5	1154.2	99.4	1171	4 AAS31128	AAS31128 Human dia
6	1148.4	98.9	1178	8 ACD03627	ACD03627 Novel hum
7	1140.4	98.2	1178	12 ADM93810	ADM93810 DNA enco
8	1120.8	96.5	1171	6 ABO54564	ABO54564 Human ova
9	1084	93.4	1211	10 ADF94944	ADF94944 Human gen
10	1082.8	93.3	1114	10 ADF94916	ADF94916 Human gen
11	988.4	85.1	993	8 ACC42468	ACC42468 Human ClQ
12	930.4	80.1	932	3 AAZ45606	AAZ45606 CDNA enco
13	703.6	60.6	885	8 ACC42469	ACC42469 Murine CI
14	702.2	60.5	1107	3 AAZ61744	AAZ61744 CDNA enco
15	702.2	60.5	1107	4 AAC99677	AAC99677 Skin cell
16	702.2	60.5	1107	6 ABL34829	ABL34829 Rat CDNA
17	701.6	60.4	1052	3 AAZ61811	AAZ61811 CDNA enco
18	701.6	60.4	1052	4 AAC99744	AAC99744 Skin cell
19	695	59.9	855	3 ABL34896	ABL34896 Rat CDNA
20	627.4	54.0	886	3 AAC67788	AAC67788 Murine AC
21	627.4	54.0	886	3 AAZ45688	AAZ45688 CDNA enco

22	627.4	54.0	886	8 AB224308	AB224308 Human GMG
23	613.6	52.9	855	3 AAC65137	AAC65137 Adipocyte
24	613.6	52.9	855	11 ADM66430	ADM66430 Degenerat
25	613.6	52.9	855	12 ADL26865	ADL26865 Human zac
26	507	43.7	855	3 AAZ45689	AAZ45689 EST with
27	507	43.7	855	3 AAZ45607	AAZ45607 CDNA enco
28	453	39.0	750	4 AAZ61624	AAZ61624 CDNA enco
29	453	39.0	750	4 AAC99557	AAC99557 Skin cell
30	453	39.0	750	6 ABL34709	ABL34709 Rat CDNA
31	337.4	29.1	909	4 AAC89872	AAC89872 Human zac
32	309	26.6	1282	4 AAC89875	AAC89875 Mouse zac
33	295.4	25.4	870	4 AAD16350	AAD16350 Human SBH
34	295.4	25.4	870	6 ABO86182	ABO86182 Novel hum
35	295.4	25.4	870	10 ADF82821	ADF82821 Neurologi
36	295.4	25.4	912	4 AAC89867	AAC89867 Human zac
37	295.4	25.4	912	4 AAD16351	AAD16351 Human SBH
38	295.4	25.4	1242	6 AAL44067	AAL44067 Human gen
39	295.4	25.4	1297	6 ABR94966	ABR94966 Human nov
40	295.4	25.4	1297	10 ADF82829	ADF82829 Adiponec
41	295.4	25.4	1304	8 ADA40160	ADA40160 Human sec
42	295.4	25.4	1304	10 ADC73711	ADC73711 Human sec
43	295.4	25.4	1304	10 ADD37685	Add37685 Gene sec
44	295.4	25.4	1304	10 ADA56322	ADA56322 Gene enco
45	295.4	25.4	1336	10 ADF94915	ADF94915 Human gen

## ALIGNMENTS

RESULT 1  
ID AAC65132 standard; CDNA, 1161 BP.  
XX  
XX AAC65132;  
XX  
XX 12-FEB-2001 (first entry)  
XX  
XX Human adipocyte complement related protein homologue zacr2p2 CDNA.  
XX  
XX Human; zacr2p2; adipocyte complement related protein; Acrp30;  
XX KW energy balance; metabolism; haemostasis; anti-microbial; ss.  
XX OS Homo sapiens.  
XX  
XX WO200063376-A1.  
XX  
XX 26-OCT-2000.  
XX  
XX 19-APR-2000; 2000WO-US010452.  
XX  
XX 20-APR-1999; 99US-00295072.  
XX  
XX (ZYMO ) ZYMOGENETICS INC.  
XX  
XX Piddington CS, Bishop PD;  
XX  
XX WPI; 2000-647517/62.  
XX P-PSDB; AAB30232.  
XX  
XX Human DNA sequence encoding a zacr2p2 polypeptide which has homology to an  
XX PT adipocyte complement related protein (Acrp30), useful in gene therapy  
XX applications for inhibiting or increasing zacr2p2 activity.  
XX  
XX Claim 31; Page 110-112; 125pp; English.  
XX  
XX The present invention is related to the isolation and uses of a homologue  
XX to the adipocyte complement related protein Acrp30, known as zacr2p2. The  
XX zacr2p2 protein is involved in energy balance, and the protein, its  
XX antibodies and coding sequence can be used to modulate energy balance,  
XX haemostasis, calcium ion concentration, muscle contraction, hormone  
XX secretion, DNA synthesis and cell growth, inositol phosphate turnover,  
XX archidonate release, phospholipase-C activation, gastric emptying, human  
XX neutrophil activation or ADCC capability and superoxide anion production.

CC They may also have uses in antimicrobial applications  
 XX Sequence 1161 BP; 237 A; 375 C; 350 G; 199 T; 0 U; 0 Other;

Query Match 100.0%; Score 1161; DB 3; Length 1161;  
 Best Local Similarity 100.0%; Pred. No. 7e-263;  
 Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GGAATAATGCTGCGGGGCGAGCGCTGCGCGGCTGCGCGTGAAGGAAAGCCGGAGC 60
Db 1 GGAATAATGCTGCGGGGCGAGCGCTGCGCGGCTGCGCGTGAAGGAAAGCCGGAGC 60
Qy 61 GCGGAGCCCCGCGCGAGAGCTTCTTGTCTCGGAGCGCCCTGGAAGTGGCGGCGG 120
Db 61 GCGGAGCCCCGCGCGAGAGCTTCTTGTCTCGGAGCGCCCTGGAAGTGGCGGCGG 120
Qy 121 AGGGTAACCAACATGATCCCTGAGGTGCTCTGAGCTGTGCTCCCTGTGCTGAC 180
Db 121 AGGGTAACCAACATGATCCCTGAGGTGCTCTGAGCTGTGCTCCCTGTGCTGAC 180
Qy 181 CCACTGCTGGCGGCTTGTGCTCGAGGGAGCTTCGGAAGGCTCCCTCAACTGCTGAC 240
Db 181 CCACTGCTGGCGGCTTGTGCTCGAGGGAGCTTCGGAAGGCTCCCTCAACTGCTGAC 240
Qy 241 AGCTGCTGCGCCCGCGAGGCGCCAGGCGCCCGCGAGGCGCCCTCAGAGATG 300
Db 241 AGCTGCTGCGCCCGCGAGGCGCCAGGCGCCCGCGAGGCGCCCTCAGAGATG 300
Qy 301 ATGGGAAGATGGGCTTCTCTGCGAAGACGCGCCAAATGAGACAGCGGAGACCGGGG 360
Db 301 ATGGGAAGATGGGCTTCTCTGCGAAGACGCGCCAAATGAGACAGCGGAGACCGGGG 360
Qy 361 GACAGCGGAGGAGGAGGTCCACTGCGCGGAGAGTAAACGGGGAAGAGCAGACCAAG 420
Db 361 GACAGCGGAGGAGGAGGTCCACTGCGCGGAGAGTAAACGGGGAAGAGCAGACCAAG 420
Qy 421 GCGAAGACCGGGGCGCATTTGGGGGCTGGCCCGCTGGCGCCCAAGGGGCTCAACGATAC 480
Db 421 GCGAAGACCGGGGCGCATTTGGGGGCTGGCCCGCTGGCGCCCAAGGGGCTCAACGATAC 480
Qy 481 CCGCGGAAGAGTGGCAACCGAGGAGGAGGCGCCCAAGGGGAGGAGGAGGAGGAGGAGG 540
Db 481 CCGCGGAAGAGTGGCAACCGAGGAGGAGGCGCCCAAGGGGAGGAGGAGGAGGAGGAGG 540
Qy 541 CTCGCAAGGCGCCCTGAGCTGTGCGAGTGGCCATCAAGTCAAGTCTTCTCGGTGGAGTG 600
Db 541 CTCGCAAGGCGCCCTGAGCTGTGCGAGTGGCCATCAAGTCAAGTCTTCTCGGTGGAGTG 600
Qy 601 ACCAAGAGTACCCAGCGGAGCGGCTGCCATCAAGTTTGAACAAGATTCTGATGAACGAG 660
Db 601 ACCAAGAGTACCCAGCGGAGCGGCTGCCATCAAGTTTGAACAAGATTCTGATGAACGAG 660
Qy 661 GGTGGCCATCAATGCTTCCAGGGGCAAGTTCCTTGGCGGCTGGCTGGGATCTATCAC 720
Db 661 GGTGGCCATCAATGCTTCCAGGGGCAAGTTCCTTGGCGGCTGGCTGGGATCTATCAC 720
Qy 721 TTCACTACGACATCAAGCTGTGGCAACAAGCACTGGGCATCGGCTGTGTCACAACGAGC 780
Db 721 TTCACTACGACATCAAGCTGTGGCAACAAGCACTGGGCATCGGCTGTGTCACAACGAGC 780
Qy 781 CAGTACCGCATCCGAGCTTTTGTATGCAACAACCGGCAACAACGATGTGGCTCAAGGCTCC 840
Db 781 CAGTACCGCATCCGAGCTTTTGTATGCAACAACCGGCAACAACGATGTGGCTCAAGGCTCC 840
Qy 841 ACCATCTGCTCTCAAGCAAGGAGTGAAGAAATTTGGCTGCAAGATCTTCTACTCAGAGCAG 900
Db 841 ACCATCTGCTCTCAAGCAAGGAGTGAAGAAATTTGGCTGCAAGATCTTCTACTCAGAGCAG 900
Qy 901 AACGGGCTCTTCTATGACCTTTATGACAGACAGCCTTTTAAAGGGCTTCTTATATCTAT 960
Db 901 AACGGGCTCTTCTATGACCTTTATGACAGACAGCCTTTTAAAGGGCTTCTTATATCTAT 960
Qy 961 GCCGACAGATGACCCCAACGAGGTATGACATGCCAGCGGCTCTCAAGGACGAGGAA 1020
  
```

```

Db 961 GCCGACAGATGACCCCAACGAGGTATGACATGCCAGCGGCTCTCAAGGACGAGGAA 1020
Qy 1021 CAAGCTTCTGACCTTGGGCTTACAGAGCAAGACCCCAACTGATGAGCTGGGGTGGGG 1080
Db 1021 CAAGCTTCTGACCTTGGGCTTACAGAGCAAGACCCCAACTGATGAGCTGGGGTGGGG 1080
Qy 1081 GTCAGTGAAGGCTTTCAGGCTCAGGCTCAAGCTCTCGGCTCTTTTTCCTTCAAT 1140
Db 1081 GTCAGTGAAGGCTTTCAGGCTCAGGCTCAAGCTCTCGGCTCTTTTTCCTTCAAT 1140
Qy 1141 AATTCAAACCTTTTATTC 1161
Db 1141 AATTCAAACCTTTTATTC 1161

RESULT 2
AB24307
ID AB24307 standard; DNA; 1161 BP.
XX
AC AB24307;
XX
DT 15-MAY-2003 (first entry)
XX
DE Human GMG-1 polypeptide encoding DNA.
XX
KW GMG-1; fatty acid; anorectic; antidiabetic; antiarteriosclerotic;
KW hypotensive; cerebroprotective; cardiatic; antilipemic; immunomodulator;
KW gene therapy; human; gene; de.
XX
OS Homo sapiens.
XX
FH Key 1.132 Location/Qualifiers
FT 5'UTR /*tag= a
FT CDS 133..990 /*tag= b
FT /*product= "GMG-1"
FT 3'UTR 991..1161 /*tag= c
FT
XX
PN W0203010197-A2.
XX
PD 06-FEB-2003.
XX
PE 23-JUL-2002; 2002W0-1B003402.
XX
PR 25-JUL-2001; 2001US-0307876P.
PR 30-JUL-2001; 2001US-0308979P.
PR 31-JUL-2001; 2001US-0309197P.
PR 23-AUG-2001; 2001US-0314575P.
XX
PA (GENST ) GENSET SA.
XX
XX
XX Lucas J, Scalia A, Dialynas D;
XX
XX MPI; 2003-229638/22.
XX
XX P-PSDB; ABP71425.
XX
XX Lowering circulating free fatty acid levels in an individual using GMG-1
XX polypeptides and polynucleotides, useful in treating obesity and obesity-
XX related diseases, e.g. diabetes, atherosclerosis, hypertension and
XX strokes.
XX
XX Disclosure; Page 126-127; 130pp; English.
XX
XX The invention relates to lowering circulating free fatty acid levels in
XX an individual. The method involves administering a composition comprising
XX a carrier and a homotrimeric GMG-1 polypeptide fragment having amino
XX acid residues 16-134 to 285 of the sequence ABP71425 or residues 16-88 to
XX 217 of the sequence ABP71426 where the cysteine at position 36 is
XX substituted with serine. The methods and compositions of the invention
XX are useful for the diagnosis and treatment of disorders associated with
  
```

CC the aberrant expression or activity of the gGmG-1 polypeptide, such as  
CC obesity and obesity-related diseases, like diabetes, atherosclerosis,  
CC hypertension, stroke, coronary insufficiency, hyperlipidemia,  
CC hypertriglyceridemia, cachexia and bulimia. The present sequence represents a  
CC DNA encoding the human gGmG-1 polypeptide

XX Sequence 1161 BP; 237 A; 375 C; 350 G; 199 T; 0 U; 0 Other;

Query Match 100.0%; Score 1161; DB 8; Length 1161;  
Best Local Similarity 100.0%; Pred. No. 7e-263;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGAAGATGATGCTGGGGGCGAGCGCTGCGCGGCTGCGCTGCGTGAAGAAAGCGGGAC 60
DB 1 GGAAGATGATGCTGGGGGCGAGCGCTGCGCGGCTGCGCTGCGTGAAGAAAGCGGGAC 60
QY 61 GCGAGGCGCGCGGAGAGCTTCTTGTGCTGCGAGCGCGCTGCGTGAAGCGGCGCG 120
DB 61 GCGAGGCGCGCGGAGAGCTTCTTGTGCTGCGAGCGCGCTGCGTGAAGCGGCGCG 120
QY 121 AGGGTAACCAACATGATCCCTGCGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGAC 180
DB 121 AGGGTAACCAACATGATCCCTGCGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGAC 180
QY 181 CCACTGCTTGGCGCTTGTGCTGCGAGGGAATTCGAGAAAGCTCCCTCAACTGCTGCTGC 240
DB 181 CCACTGCTTGGCGCTTGTGCTGCGAGGGAATTCGAGAAAGCTCCCTCAACTGCTGCTGC 240
QY 241 AGCCTGCTGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAGAGATG 300
DB 241 AGCCTGCTGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAGAGATG 300
QY 301 ATGGAGAGATGAGGCTTCTTCTGCGAAAGCGCGCAAGATGACAGAGCGCGCGCG 360
DB 301 ATGGAGAGATGAGGCTTCTTCTGCGAAAGCGCGCAAGATGACAGAGCGCGCGCG 360
QY 361 GACAGCGAGAGAGAGAGGTCACCTGCGCGGACAGTAAACGGGAGAAAGCGAGCAAG 420
DB 361 GACAGCGAGAGAGAGAGGTCACCTGCGCGGACAGTAAACGGGAGAAAGCGAGCAAG 420
QY 421 GGGAGAGCGCGCGCGCATTTGGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCTCA 480
DB 421 GGGAGAGCGCGCGCGCATTTGGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCTCA 480
QY 481 CCGCGAGAGATGCGACACAGCGAGCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 CCGCGAGAGATGCGACACAGCGAGCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 CTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 CTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 ACCAAGAGCTTACCGAGCGGAGCGGCTGCCATCAAGTTTGAACAAGTTTGAATGA 660
DB 601 ACCAAGAGCTTACCGAGCGGAGCGGCTGCCATCAAGTTTGAACAAGTTTGAATGA 660
QY 661 GGTGGCGCACTAATGCTTCAGCGGCAAGTTGCTGCGCGCGCTGCGGATCTACTAC 720
DB 661 GGTGGCGCACTAATGCTTCAGCGGCAAGTTGCTGCGCGCGCTGCGGATCTACTAC 720
QY 721 TTGACCTAGAGCATACGCTGCGCAAGCAAGCTTGGCGCGCTGCGGATCTACTAC 780
DB 721 TTGACCTAGAGCATACGCTGCGCAAGCAAGCTTGGCGCGCTGCGGATCTACTAC 780
QY 781 CAGTACCGATCCGAGCTTGTGATGCAACCGCGCAACAGATGAGCTTCAAGGCTCC 840
DB 781 CAGTACCGATCCGAGCTTGTGATGCAACCGCGCAACAGATGAGCTTCAAGGCTCC 840
QY 841 ACCATCTGCTCTCAAGCGGAGGAGAAATTGCTGCAAGATCTTCTACTCAGAGCAG 900
DB 841 ACCATCTGCTCTCAAGCGGAGGAGAAATTGCTGCAAGATCTTCTACTCAGAGCAG 900
QY 901 AACGGGCTTCTATGACCTTACTGAGACAGACGCTTCTTACGGGCTTCTAATCTAT 960

```

```

DB 901 AACGGGCTTCTATGACCTTACTGAGACAGACGCTTCTTACGGGCTTCTAATCTAT 960
QY 961 GCCGACGAGATGACCCCAACGAGATATGACATGCCAGCGGCTCTCCAGCGAGGAA 1020
DB 961 GCCGACGAGATGACCCCAACGAGATATGACATGCCAGCGGCTCTCCAGCGAGGAA 1020
QY 1021 CAAGCTTCTGAGCTTGGGCTTACGAGCAAGACCCCAACTGTAGAGCTGGGGGTGGGG 1080
DB 1021 CAAGCTTCTGAGCTTGGGCTTACGAGCAAGACCCCAACTGTAGAGCTGGGGGTGGGG 1080
QY 1081 GTGAGTGAAGGCTTACGCTCAGGCTCAGCTCAGCTCAGCTCTTTTTCCTTCAAT 1140
DB 1081 GTGAGTGAAGGCTTACGCTCAGGCTCAGCTCAGCTCAGCTCTTTTTCCTTCAAT 1140
QY 1141 AATTCAAACCTTTTATTC 1161
DB 1141 AATTCAAACCTTTTATTC 1161

```

# RESULT 3

```

ADME6421
ID ADM6421 standard; CDNA; 1161 BP.
XX
AC ADM6421;
XX
DT 03-JUN-2004 (first entry)
XX
DE cDNA encoding human ACRP homologue, zacr2.
XX
KW Human; adipocyte complement related protein; ACRP; zacr2;
KW collagen-like domain; C1q domain; inflammatory disorder;
KW microbial infection; wound healing; antiinflammatory; vulnary;
KW antimicrobial; chromosome 5q34; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 133..990
FT FT /*tag= a
FT FT /product= "Zacr2"
XX
XX US6620909-B1.
XX
XX 16-SEP-2003.
XX
XX 19-APR-2000; 2000US-00552204.
XX
XX 20-APR-1999; 99US-0130207P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Piddington CS, Bishop PD;
XX
XX WPI: 2003-895428/82.
XX
XX P-PSDB; ADM6422.
XX
XX New zacr2 polypeptide, useful for preparing a composition for treating
XX inflammatory disorders or microbial infections or for promoting wound
XX healing.
XX
XX Disclosure; SEQ ID NO 1; 36pp; English.
XX
XX The present invention relates to the isolation of an adipocyte complement
XX related protein (ACRP) homologue designated zacr2, and the
XX polynucleotide sequence encoding it. Zacr2 is a novel member of the
XX protein family comprising a collagen-like and a C1q domain. The zacr2
XX polypeptide is useful for preparing a composition for treating
XX inflammatory disorders, and microbial infections, or for promoting wound
XX healing. The present sequence encodes human zacr2. The gene encoding
XX human zacr2 maps to chromosome 5q34.
XX
XX Sequence 1161 BP; 237 A; 375 C; 350 G; 199 T; 0 U; 0 Other;

```

Query Match 100.0%; Score 1161; DB 11; Length 1161;  
 Best Local Similarity 100.0%; Pred. No. 7e-263;  
 Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GAAATAATATGCTTGGGCGGACGCTCTGCGGCTGCGCTGAGAGAAAGCCGGAGC 60
DB 1 GAAATAATATGCTTGGGCGGACGCTCTGCGGCTGCGCTGAGAGAAAGCCGGAGC 60
QY 61 GCGGAGCCCCCGGAGAGCTTCTTGTCCGGAAGCCCCCTGAGAGTGGCGGGCGGCG 120
DB 61 GCGGAGCCCCCGGAGAGCTTCTTGTCCGGAAGCCCCCTGAGAGTGGCGGGCGGCG 120
QY 121 AGGTAACACACATGATCCCTGAGGTGCTCTGAGCTGTGCCCCCTGTGTGTGAC 180
DB 121 AGGTAACACACATGATCCCTGAGGTGCTCTGAGCTGTGCCCCCTGTGTGTGAC 180
QY 181 CCACTGCTTGGCGCTTGTGCTCGAGGAGCTTCCGGAAGGCTCCCTCAACTGCTGTC 240
DB 181 CCACTGCTTGGCGCTTGTGCTCGAGGAGCTTCCGGAAGGCTCCCTCAACTGCTGTC 240
QY 241 AGCTGCTTGGCGCTTGTGCTCGAGGAGCTTCCGGAAGGCTCCCTCAACTGCTGTC 300
DB 241 AGCTGCTTGGCGCTTGTGCTCGAGGAGCTTCCGGAAGGCTCCCTCAACTGCTGTC 300
QY 301 ATGGGAGCAATGGGCTTCTGCGAAGACGCGCAAGTGAACAGCGGAGCCGGAGG 360
DB 301 ATGGGAGCAATGGGCTTCTGCGAAGACGCGCAAGTGAACAGCGGAGCCGGAGG 360
QY 361 GACAGCGGAGAGGAGGTCCAAGTGTGCGGACAGTAAACGGGAGAAAGCCAGAAC 420
DB 361 GACAGCGGAGAGGAGGTCCAAGTGTGCGGACAGTAAACGGGAGAAAGCCAGAAC 420
QY 421 GCGAAGACCGGGGCGCTTGGCGGGCTGCGCCCCCGTGGCCCAAGGGGCTCAAGGTACC 480
DB 421 GCGAAGACCGGGGCGCTTGGCGGGCTGCGCCCCCGTGGCCCAAGGGGCTCAAGGTACC 480
QY 481 CCGGAGAGAGTGGCACACAGGCAAGAGGGGCCCAAGGCAAGAAAGGGAGCCAGGC 540
DB 481 CCGGAGAGAGTGGCACACAGGCAAGAGGGGCCCAAGGCAAGAAAGGGAGCCAGGC 540
QY 541 CTCCCAAGGCGCTTGCAGCTGTGGCAGTGGCCATCAAGTCAAGCTTTCTCGTGGCAGTG 600
DB 541 CTCCCAAGGCGCTTGCAGCTGTGGCAGTGGCCATCAAGTCAAGCTTTCTCGTGGCAGTG 600
QY 601 ACCAAGAGCTTACCAAGGAGCGGCTGCCATCAAGTTTGAAGAGTTTGAATGAAGAG 660
DB 601 ACCAAGAGCTTACCAAGGAGCGGCTGCCATCAAGTTTGAAGAGTTTGAATGAAGAG 660
QY 661 GGTGGCACTTCAATGCTTCCAGCGGCAAGTTGCTGCGGCGTGGCTGGATCTACTAC 720
DB 661 GGTGGCACTTCAATGCTTCCAGCGGCAAGTTGCTGCGGCGTGGCTGGATCTACTAC 720
QY 721 TTCACTTGAAGATCAGCTGCTGCAAGAGCACTGCGCATGGCTGTGTGACACAGGC 780
DB 721 TTCACTTGAAGATCAGCTGCTGCAAGAGCACTGCGCATGGCTGTGTGACACAGGC 780
QY 781 CAGTACCGCAATCGGAGCTTGTGATGCAACACCGGCAACCAAGTGTGGCTTCAGGCTCC 840
DB 781 CAGTACCGCAATCGGAGCTTGTGATGCAACACCGGCAACCAAGTGTGGCTTCAGGCTCC 840
QY 841 ACCATCTGAGCTCTCAAGAGGGTGAAGAGTTGGCTGCGAGATCTTCTACTCAGACAG 900
DB 841 ACCATCTGAGCTCTCAAGAGGGTGAAGAGTTGGCTGCGAGATCTTCTACTCAGACAG 900
QY 901 AACGGGCTCTTATATGACCTTACTGAGACAGACCTCTTTTACGGGCTTCTAATCTAT 960
DB 901 AACGGGCTCTTATATGACCTTACTGAGACAGACCTCTTTTACGGGCTTCTAATCTAT 960
QY 961 GCGGACAGAGATGACCCCAAGAGAGTATAGCATGCAAGCGGAGTCTTCCAGGAGGAAA 1020
DB 961 GCGGACAGAGATGACCCCAAGAGAGTATAGCATGCAAGCGGAGTCTTCCAGGAGGAAA 1020

```

```

QY 1021 CAAGCTTCTGACTTGGGCTTAAAGAGCAAGCCCAACACTGATGCTGGGCTGGGG 1080
DB 1021 CAAGCTTCTGACTTGGGCTTAAAGAGCAAGCCCAACACTGATGCTGGGCTGGGG 1080
QY 1081 GTGAGTGAAGCGGTTCTAGGCTTCAAGGCTCAACTCTCGGCTCTTTTTCCTTCAAT 1140
DB 1081 GTGAGTGAAGCGGTTCTAGGCTTCAAGGCTCAACTCTCGGCTCTTTTTCCTTCAAT 1140
QY 1141 AAATCCAAACCTTTTATTTCA 1161
DB 1141 AAATCCAAACCTTTTATTTCA 1161

```

## RESULT 4

ADL26856 standard; DNA, 1161 BP.

ADL26856;

20-MAY-2004 (first entry)

Human zacr2p DNA.

Adipocyte complement related protein homologue; zacr2p; energy balance;  
 energy efficiency; infection; cellular metabolism; tumour;  
 cardiopulmonary bypass ischaemia; recesitation; myocardial infarction;  
 post-trauma vasospasm; stroke; percutaneous transluminal angioplasty;  
 vasculature disease; disseminated intravascular coagulation;  
 non-vascular disease; arteriosclerosis; prosthetic biomaterial;  
 surgical equipment; wound healing; gene therapy; human; chromosome 5;  
 gene; ds.

Homo sapiens.

Location/Qualifiers

133..990

/\*tag= b

/product= "zacr2p protein"

133..177

/\*tag= a

178..987

/\*tag= c

/product= "Mature zacr2p protein"

US2004024187-A1.

05-FEB-2004.

17-JUL-2003; 2003US-00621787.

20-APR-1999; 99US-0130207P.

19-APR-2000; 2000US-00552204.

(ZYMO) ZYMOGENETICS INC.

Piddington CS, Bishop PD;

WPI; 2004-142676/14.

P-PSDB; ADL26857.

New zacr2p polypeptide, useful in modulating energy balance, treating

tumors, myocardial infarction, stroke, disseminated intravascular

coagulation or arteriosclerosis, mediating wound repair and healing or in

enhancing antibody formation.

Claim 31; SEQ ID NO 1; 41bp; English.

The invention relates to novel adipocyte complement related protein

homologue, zacr2p polypeptides and polymucleotides. Sequences of the

invention are useful in modulating energy balance, analysing energy

efficiency of a mammal and in studying insulin, exogenous microorganism

infection, cellular metabolism, dimerisation or oligomerisation. They are

also useful in treating tumours, cardiopulmonary bypass ischaemia and



CC resection, myocardial infarction and post-trauma vasospasm e.g. stroke  
 CC or percutaneous transluminal angioplasty, vasculature disease e.g.  
 CC disseminated intravascular coagulation and non-vascular disease e.g.  
 CC arteriosclerosis. zacr2 sequences are also useful in pacifying damaged  
 CC collagenous tissue, prosthetic biomaterials and surgical equipments,  
 CC preventing infection, mediating wound repair or wound healing and in  
 CC enhancing antibody formation. The invention is also used in gene therapy.  
 CC The present sequence is human zacr2 DNA. zacr2 gene is located on  
 CC chromosome 5.

XX Sequence 1161 BP, 237 A; 375 C; 350 G; 199 T; 0 U; 0 Other;

Query Match 100.0%; Score 1161; DB 12; Length 1161;

Best Local Similarity 100.0%; Pred. No. 7e-263; Mismatches 0; Gaps 0;

Matches 1161; Conservative 0; Indels 0; Gaps 0;

QY 1 GGAATACTATGCTGGGCGGACGCTGCGCGCTGCGCTGCGCTGAGAAAGCGGAGC 60  
 DB 1 GGAATACTATGCTGGGCGGACGCTGCGCGCTGCGCTGAGAAAGCGGAGC 60

QY 61 GCGAGGCGCGCGCGAGCTTCTTCTGCGGAGCGCGCTGAGAGTGGCGGCGCGG 120  
 DB 61 GCGAGGCGCGCGCGAGCTTCTTCTGCGGAGCGCGCTGAGAGTGGCGGCGCGG 120

QY 121 AGGATAACACAGATCCCTGGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGAC 180  
 DB 121 AGGATAACACAGATCCCTGGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGAC 180

QY 181 CCACTGCTGGCGCTTGTCTGCGAGGAGCTTCCGGAAGGCTCCCTCACTGCTGTC 240  
 DB 181 CCACTGCTGGCGCTTGTCTGCGAGGAGCTTCCGGAAGGCTCCCTCACTGCTGTC 240

QY 241 AGCTGCTGCG 300  
 DB 241 AGCTGCTGCG 300

QY 301 ATGGGAGGAGTGGGCTTCTGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
 DB 301 ATGGGAGGAGTGGGCTTCTGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 360

QY 361 GACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
 DB 361 GACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420

QY 421 GCGAAGCG 480  
 DB 421 GCGAAGCG 480

QY 481 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 DB 481 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

QY 541 CTCGAGGCG 600  
 DB 541 CTCGAGGCG 600

QY 601 ACCAAGGCTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
 DB 601 ACCAAGGCTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660

QY 661 GGTGGGCACTACATGCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
 DB 661 GGTGGGCACTACATGCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

QY 721 TTCACTACGACATCACTGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
 DB 721 TTCACTACGACATCACTGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780

QY 781 CAGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 DB 781 CAGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840

QY 841 ACGATCTGGCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
 DB 841 ACGATCTGGCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900

DB 841 ACGATCTGGCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900

QY 901 AACGGGCTCTTATGACCTTACTGAGACAGACAGCTCTTTACGGGCTTCTATCTAT 960  
 DB 901 AACGGGCTCTTATGACCTTACTGAGACAGACAGCTCTTTACGGGCTTCTATCTAT 960

QY 961 GCGAGCAGGATGACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
 DB 961 GCGAGCAGGATGACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020

QY 1021 CAAGCTTCTGAGCTTGGGCTTAAAGAGCAAGACCCCAACTGATGAGGCTGGGGG 1080  
 DB 1021 CAAGCTTCTGAGCTTGGGCTTAAAGAGCAAGACCCCAACTGATGAGGCTGGGGG 1080

QY 1081 GTGAGTGAAGCGGTTTCTAGGCTCAAGCTCTCTCGCTCTTTTCTTCTTCTAT 1140  
 DB 1081 GTGAGTGAAGCGGTTTCTAGGCTCAAGCTCTCTCTCGCTCTTTTCTTCTTCTAT 1140

QY 1141 AATCCAACTTTTATCA 1161  
 DB 1141 AATCCAACTTTTATCA 1161

RESULT 5  
 AAS31128  
 ID AAS31128 standard; cDNA; 1171 BP.  
 XX AAS31128;  
 AC AAS31128;  
 DT 04-DEC-2001 (first entry)  
 XX

DE Human diagnostic and therapeutic polynucleotide (DITHP) #143.  
 XX  
 KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
 KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
 KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
 KW respiratory disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200162927-A2.  
 XX  
 PD 30-AUG-2001.  
 XX  
 XX 21-FEB-2001; 2001WO-US006059.  
 XX  
 PR 24-FEB-2000; 2000US-0184693P.  
 PR 24-FEB-2000; 2000US-0184697P.  
 PR 24-FEB-2000; 2000US-0184688P.  
 PR 24-FEB-2000; 2000US-0184768P.  
 PR 24-FEB-2000; 2000US-0184769P.  
 PR 24-FEB-2000; 2000US-0184770P.  
 PR 24-FEB-2000; 2000US-0184771P.  
 PR 24-FEB-2000; 2000US-0184772P.  
 PR 24-FEB-2000; 2000US-0184773P.  
 PR 24-FEB-2000; 2000US-0184774P.  
 PR 24-FEB-2000; 2000US-0184776P.  
 PR 24-FEB-2000; 2000US-0184777P.  
 PR 24-FEB-2000; 2000US-0184797P.  
 PR 24-FEB-2000; 2000US-0184813P.  
 PR 24-FEB-2000; 2000US-0184837P.  
 PR 24-FEB-2000; 2000US-0184841P.  
 PR 24-FEB-2000; 2000US-0185213P.  
 PR 24-FEB-2000; 2000US-0185216P.  
 PR 12-MAY-2000; 2000US-0203785P.  
 PR 15-MAY-2000; 2000US-0204226P.  
 PR 16-MAY-2000; 2000US-0204525P.  
 PR 16-MAY-2000; 2000US-0204821P.  
 PR 16-MAY-2000; 2000US-0204908P.  
 PR 16-MAY-2000; 2000US-0205232P.  
 PR 17-MAY-2000; 2000US-0204815P.  
 PR 17-MAY-2000; 2000US-0204863P.





Db 847 ACCATCTGCTCTCAAGCAGGGGAGAGAGTTGGCTGAGACTCTTACTAGAGAG 906  
 Qy 901 AACGGGCTCTTATAGACCTTACTGACAGACGCTCTTTACGGGCTTCTAATCTAT 960  
 Db 907 AACGGGCTCTTATAGACCTTACTGACAGACGCTCTTTACGGGCTTCTAATCTAT 966  
 Qy 961 GCGAGCAGAGATGACCCCAAGAGATATAGCATGCGACGGGCTCTCCAGGAGGAAA 1020  
 Db 967 GCGAGCAGAGATGACCCCAAGAGATATAGCATGCGACGGGCTCTCCAGGAGGAAA 1026  
 Qy 1021 CAAGCTTCTGAGCTTGGGCTTACAGACAGACCCCAACTGTAGGCTGGGGGTGGAG 1080  
 Db 1027 CAAGCTTCTGAGCTTGGGCTTACAGACAGACCCCAACTGTAGGCTGGGGGTGGAG 1086  
 Qy 1081 GTGAGTGAAGGGGTTCTAGGCTCAGGCTCAGCTCTCGGCTC-TTTTTTCCCTTCAT 1139  
 Db 1087 GTGAGTGAAGGGGTTCTAGGCTCAGGCTCAGCTCTCGGCTC-TTTTTTCCCTTCAT 1146  
 Qy 1140 TAAATCAAACTTTTATTCA 1161  
 Db 1147 TAAATCAAACTTTTATTCA 1168  
 RESULT 7  
 ADM93810  
 ID ADM93810 standard; DNA, 1178 BP.  
 AC ADM93810;  
 XX 17-JUN-2004 (first entry)  
 DE DNA encoding human NOV protein #15.  
 XX  
 KW gene therapy; vaccine; ds; gene; NOX; cancer;  
 KM neurodegenerative disorder; Parkinson's disease; metabolic disorder;  
 KM diabetes; obesity; immune related disorder; tissue typing; human.  
 OS Homo sapiens.  
 XX  
 PN US2004009480-A1.  
 PD 15-JAN-2004.  
 XX  
 PF 03-JUN-2002; 2002US-00162335.  
 XX  
 PR 04-JUN-2001; 2001US-0293607P.  
 PR 04-JUN-2001; 2001US-0293661P.  
 PR 06-JUN-2001; 2001US-0296404P.  
 PR 06-JUN-2001; 2001US-0296418P.  
 PR 11-JUN-2001; 2001US-0297414P.  
 PR 12-JUN-2001; 2001US-0297567P.  
 PR 14-JUN-2001; 2001US-0298285P.  
 PR 15-JUN-2001; 2001US-0298556P.  
 PR 21-JUN-2001; 2001US-0299949P.  
 PR 26-JUN-2001; 2001US-0300883P.  
 PR 28-JUN-2001; 2001US-0301550P.  
 PR 13-AUG-2001; 2001US-0311972P.  
 PR 27-AUG-2001; 2001US-0315069P.  
 PR 27-AUG-2001; 2001US-0315071P.  
 PR 29-AUG-2001; 2001US-0315660P.  
 PR 14-SEP-2001; 2001US-0322939P.  
 PR 17-SEP-2001; 2001US-0322706P.  
 PR 14-DEC-2001; 2001US-0341186P.  
 PR 28-FEB-2002; 2002US-0363189P.  
 PR 12-MAR-2002; 2002US-0363673P.  
 PR 12-MAR-2002; 2002US-0363676P.  
 XX  
 PA (ANDE/) ANDERSON D W.  
 PA (BAUM/) BAUMGARTNER J C.  
 PA (BOLD/) BOLDIG F L.  
 PA (CASW/) CASMAN S J.  
 PA (EDIN/) EDINGER S R.

PA (GANG/) GANGOLLI E A.  
 PA (GERL/) GERLACH V.  
 PA (GORM/) GORMAN L.  
 PA (GUOX/) GUO X S.  
 PA (HJAL/) HJALT T.  
 PA (KEKU/) KEKUDA R.  
 PA (LIIL/) LI L.  
 PA (MACD/) MACDOUGALL J R.  
 PA (MALY/) MALYANKAR U M.  
 PA (MILL/) MILLET I.  
 PA (PADI/) PADIGARU M.  
 PA (PATU/) PATURAJAN M.  
 PA (PENA/) PENNA C E A.  
 PA (RAST/) RASTELLI L.  
 PA (SHIM/) SHIMKETS R A.  
 PA (STON/) STONE D J.  
 PA (SPYT/) SPYTEK K A.  
 PA (VERN/) VERNET C A M.  
 PA (VOSS/) VOSS E Z.  
 PA (ZERN/) ZERHUSEN B D.  
 XX  
 PI Anderson DM, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR,  
 PI Gangolli EA, Gerlach V, Gorman L, Guo XS, Hjal T, Kekuda R, Li L,  
 PI MacDougall JR, Malyanakar UM, Millet I, Padigar M, Paturajan M,  
 PI Pena CE, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CM;  
 PI Voss EZ, Zerhusen BD;  
 XX  
 DR WPI; 2004-090456/09.  
 DR P-PSDB; ADM93811.  
 XX  
 PT New NOVX polypeptide, useful for preparing a composition for treating or  
 PT preventing e.g., cancer, neurodegenerative disorders such as Parkinson's  
 PT disease, or metabolic disorders such as diabetes or obesity, or for  
 PT tissue typing.  
 XX  
 PS Disclosure; SEQ ID NO 29; 202pp; English.  
 XX  
 CC The invention relates to an isolated NOVX polypeptide. The polypeptide is  
 CC useful for preparing a composition for treating or preventing a pathology  
 CC associated with NOVX polypeptide e.g. cancer, neurodegenerative disorders  
 CC such as Parkinson's disease, metabolic disorders such as diabetes or  
 CC obesity or immune related disorders or for tissue typing. The present  
 CC sequence represents DNA encoding a human NOV protein.  
 CC  
 SQ Sequence 1178 BP; 244 A; 382 C; 350 G; 202 T; 0 U; 0 Other;  
 Query Match 98.2%; Score 1140.4; DB 12; Length 1178;  
 Best Local Similarity 99.4%; Pred. No. 4.9e-258;  
 Matches 1155; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
 Qy 1 GGAAGACTATGCTGGGGGCGGAGGCTGCGCGGCTGCGCTGAGGAAAGCGGGAGC 60  
 Db 7 GGAAGACTATGCTGGGGGCGGAGGCTGCGCGGCTGCGCTGAGGAAAGCGGGAGC 66  
 Qy 61 GCGAGCCCGCGGAGGCTTCTTGGCTCCGAGCGCCCTTGAAGTGGCGGAGCGCG 120  
 Db 67 GCGAGCCCGCGGAGGCTTCTTGGCTCCGAGCGCCCTTGAAGTGGCGGAGCGCG 126  
 Qy 121 AGGTAACACATGATCCCTGAGGCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 180  
 Db 127 AGGTAACACATGATCCCTGAGGCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 186  
 Qy 181 CCATGCTTGGGCGCTTGTGCTCCAGGAGCTTCCGAAAGGCTCCCTCAAGTGTGTC 240  
 Db 187 CCATGCTTGGGCGCTTGTGCTCCAGGAGCTTCCGAAAGGCTCCCTCAAGTGTGTC 246  
 Qy 241 AGCTGCTGAGCCCGGAGGCGGAGCGGCGCCCGGAGGCGCCCGGAGGCGCTGAGATG 300  
 Db 247 AGCTGCTGAGCCCGGAGGCGGAGCGGCGCCCGGAGGCGCCCGGAGGCGCTGAGATG 306  
 Qy 301 ATGGAGCAATGGGCTTCTTGGCAAGAGCGGCAAGATGAGACGAGCGGAGCGGGGG 360  
 Db 307 ATGGAGCAATGGGCTTCTTGGCAAGAGCGGCAAGATGAGACGAGCGGAGCGGGGG 366

QY	361	CACACCGGAGAGAAAGGTCCACTCTGGCCCGGACAGGTTAACCTCGGGGAAAGCCAGACCCAAAG	420
Db	367	GACACCGGAGAGAAAGGTCCACTCTGGCCCGGACAGGTTAACCTCGGGGAAAGCCAGACCCAAAG	426
QY	421	GGCAAAAGCCGGGGCCATTGGGGGGGGGTGGCCCCCGGTGGCCCCCGGACCCCAAGGGGGGTCAACGGTTAC	480
Db	427	GGCAAAAGCCGGGGCCATTGGGGGGGGGTGGCCCCCGGTGGCCCCCGGACCCCAAGGGGGGTCAACGGTTAC	486
QY	481	CCCGGAAAGCATGGCAACACAGGACGAGAAAGGGGCCCAAGGGGCAAGAAAGGGAGCCAGGC	540
Db	487	CCCGGAAAGCATGGCAACACAGGACGAGAAAGGGGCCCAAGGGGCAAGAAAGGGAGCCAGGC	546
QY	541	CTCCAGGGCCCCCTGGACGCTGTGGCAGTGGCCATACCAAGTCAGCTTTCTCGTGGGCAAGTg	600
Db	547	CTCCAGGGCCCCCTGGACGCTGTGGCAGTGGCCATACCAAGTCAGCTTTCTCGTGGGCAAGTg	606
QY	601	ACCAAGAGCTACCCCAACGGGAGCGGCTGCCCATCAAGTTTGAACAAGATTCTGATGAACGAG	660
Db	607	ACCAAGAGCTACCCCAACGGGAGCGGCTGCCCATCAAGTTTGAACAAGATTCTGATGAACGAG	666
QY	661	GGTGGCCACTACAAATGCTTCCAGCCGCAAGTTCTGTCTGGGGCGTGCCTGGGAATCTACTAC	720
Db	667	GGTGGCCACTACAAATGCTTCCAGCCGCAAGTTCTGTCTGGGGCGTGCCTGGGAATCTACTAC	726
QY	721	TTCACTTACGACATACGCTGGGCAACAAGACCTGGCCATCGGCTGGTGCAACAACGCG	780
Db	727	TTCACTTACGACATACGCTGGGCAACAAGACCTGGCCATCGGCTGGTGCAACAACGCG	786
QY	781	CAGTACCGGCATCCGGACCTTTTGATGCAACACCCGGCAACCAAGATGTGGCTTCAGGCTCC	840
Db	787	CAGTACCGGCATCCGGACCTTTTGATGCAACACCCGGCAACCAAGATGTGGCTTCAGGCTCC	846
QY	841	ACCATCTTGGCTCTCAAGAGGGGTGACGAAGTTTGGCTGCAATCTTTAATCAAGACAG	900
Db	847	ACCATCTTGGCTCTCAAGAGGGGTGACGAAGTTTGGCTGCAATCTTTAATCAAGACAG	906
QY	901	AACGGGCTCTTCTATGACCCCTTACTGGAACAAGCGCTCTTTACGGGCTTCTTAATCTAT	960
Db	907	AACGGGCTCTTCTATGACCCCTTACTGGAACAAGCGCTCTTTACGGGCTTCTTAATCTAT	966
QY	961	GCCGACAGAGATGACCCCAACGAGGATTATGACATGCAAGCGCGGTCTTCCAGGCAAGGAA	1020
Db	967	GCCGACAGAGATGACCCCAACGAGGATTATGACATGCAAGCGCGGTCTTCCAGGCAAGGAA	1026
QY	1021	CAAGTTTCTGACCTTGGGCTTACAGAGCAAGGCCCAACATGTAGGCTGGGGGTGGGG	1080
Db	1027	CAAGTTTCTGACCTTGGGCTTACAGAGCAAGGCCCAACATGTAGGCTGGGGGTGGGG	1086
QY	1081	GTCGAGTAGCGGTTCTAGCCTCAAGGCTCAAGCTCTCCGCGCTC - TTTTTCCTCCCTTCAT	113
Db	1087	GTCGAGTAGCGGTTCTAGCCTCAAGGCTCAAGCTCTCCGCGCTCTTTTTCCTCCCTTCAT	114
QY	1140	TAAATCCAAACTTTTATTTCA 1161	
Db	1147	TAAATCCAAACTTTTATTTCA 1168	

RESULT	8
ABQ54564	
ID	ABQ54564 standard; cDNA; 1171 BP.

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HDHBB13 cDNA, SEQ ID NO:444.

KM Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KM ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KM PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
KM inflammatory condition; immune disorder; blood disorder;

KM cardiovascular disorder; respiratory disorder; neurological disorder  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KM gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytotoxic; immunomodulatory; neuroprotective;  
 KM antiinflammatory; gynaecological; reproductive; gene; ss.

**Homo sapiens.**

WO200200677-A1.

03-JAN-2002.

07-JUN-2001; 2001WO-US018569.

07-JUN-2000; 2000US-0209467P.

(HUMA-) HUMAN GENOME SCI INC.

Birba CE, Rosen CA;

WPI: 2002-147878/19.

1

useful in the prevention

diseases.

Claim 1; SEQ ID NO 4

The invention relates

encompasses polypept

recombinant vectors

of ovarian antigen F

disorders. Such conc

disorders (e.g., inf

disorders, infection

vaginitis), immune d

blood-related disorders

and urinary system c

modulate ovarian ant

### Identification of Ir

useful in disease di

invention. Note: The

from WIPO at ftp.wipo

Sequence 1171 BP; 25

### Every Match

tches 1137; Conserva

18 GCCGACGCTCT

1 GCCGACGCTCT

76 GAGCTTCTTTG  
|||||

61 GAGCTTCTTTG

Query Match 96.5%; Score 1120.8; DB 6; Length 1171;  
 Best Local Similarity 99.2%; Pred. No. 28-53;  
 Matches 1137; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy	18	GCGGACGCTCGCCGCGCTCGTGGCGGTGAGAAAGCGCGGAGCGCC--CGCGA	75
Dy	1	GCCACGCTCTGCGCCGCTGCTGCGGTGAGAAAGCGGAGCGCCGCGCGACGAG	60
Qy	76	GAGCTTCCTTGGCTCGGACGCGCCCTGAGCGTGGCGGAGCGCGAGGTTACCA	135
Dy	61	GAGCTTCCTTGGCTCGGACGCGCCCTGAGCGTGGCGGAGCGCGAGGTTACCA	120



Best Local Similarity 99.8%; Pred. No. 8.8e-245;  
Matches 1084; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 76 GAGCTTCTTTGCTCCGAGACGCCCCCTGACGTCGCGGCGAGCCGCGAGGGTAACCAACATG 135
DB 1111 GAGCTTCTTTGCTCCGAGACGCCCCCTGACGTCGCGGCGAGCCGCGAGGGTAACCAACATG 1052
QY 136 ATCCCCGGGGTCTCTCGGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTG 195
DB 1051 ATCCCCGGGGTCTCTCGGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTG 992
QY 196 TTTGCTCGAGGAGACTTCGAGAAAGGCTCCCTCAATGTCGTCGAGCTCTGAGCCCC 255
DB 991 TTTGCTCGAGGAGACTTCGAGAAAGGCTCCCTCAATGTCGTCGAGCTCTGAGCCCC 932
QY 256 CAGGCCCCACCCGCCCCCGGAGAGCCCTCAGAGATGATGAGAGAGAGG 315
DB 931 CAGGCCCCACCCGCCCCCGGAGAGCCCTCAGAGATGATGAGAGAGAGG 872
QY 316 TTTTCTTGCAAAAGACGGCCCAAGATGACACGAGCCGACCGGGGGGACAGCGAGAGAGAA 375
DB 871 TTTTCTTGCAAAAGACGGCCCAAGATGACACGAGCCGACCGGGGGGACAGCGAGAGAGAA 812
QY 376 GGTCCACTGGGCGGACAGGTAAACGGGAAAGCCAGGACCAAGAGGCAAGCCGGGGCC 435
DB 811 GGTCCACTGGGCGGACAGGTAAACGGGAAAGCCAGGACCAAGAGGCAAGCCGGGGCC 752
QY 436 ATTGGGCGGGCTGGCCCCCGTGCCGCGGAGGAGTCAACGGTACCCCGGGAGACATGGC 495
DB 751 ATTGGGCGGGCTGGCCCCCGTGCCGCGGAGGAGTCAACGGTACCCCGGGAGACATGGC 692
QY 496 ACACCAAGGCAAGAGGGGGCCCAAGGSCAAGAAAGGGAGCCAGGCTTCCAGGGCCCTGCG 555
DB 691 ACACCAAGGCAAGAGGGGGCCCAAGGSCAAGAAAGGGAGCCAGGCTTCCAGGGCCCTGCG 632
QY 556 AGCTGTGGAGTGGCCATACCAATGCTTCTCGGTGAGGAGTCAAGAGGCTACCA 615
DB 631 AGCTGTGGAGTGGCCATACCAATGCTTCTCGGTGAGGAGTCAAGAGGCTACCA 572
QY 616 CGGAGCGGCTGCCCATCAAGTTTGACAAATTTGATGAAAGAGGGTGGCCACTACAT 675
DB 571 CGGAGCGGCTGCCCATCAAGTTTGACAAATTTGATGAAAGAGGGTGGCCACTACAT 512
QY 676 GCTTCCAGCGGCAAGTTGCTGTCGCGGCTGCTGGAGTCTTACTTCACTTACAGACATC 735
DB 511 GCTTCCAGCGGCAAGTTGCTGTCGCGGCTGCTGGAGTCTTACTTCACTTACAGACATC 452
QY 736 ACGGTGACCAAGACACTGGGCTGGTGGACCAAGGCGACAGTACCGGATCGG 795
DB 451 ACGGTGACCAAGACACTGGGCTGGTGGACCAAGGCGACAGTACCGGATCGG 392
QY 796 ACCTTTGATGCAACACCGGCAACCAAGATGTCCTCAGGCTTCACCATCTGGCTTC 855
DB 391 ACCTTTGATGCAACACCGGCAACCAAGATGTCCTCAGGCTTCACCATCTGGCTTC 332
QY 856 AAGGAGGGTGAAGAGTTGGCTGCAAGTCTTACTCAAGGCAAGCGGCTCTTCTAT 915
DB 331 AAGGAGGGTGAAGAGTTGGCTGCAAGTCTTACTCAAGGCAAGCGGCTCTTCTAT 272
QY 916 GACCTTACTGAGACAGACGCTCTTTAAGGGCTTCCAACTTATGCGACAGAGATGAC 975
DB 271 GACCTTACTGAGACAGACGCTCTTTAAGGGCTTCCAACTTATGCGACAGAGATGAC 212
QY 976 CCAACAGAGTATAGACATGCAACGCGGCTCTCAGGCGAGGAAACAAGCTTCTGACTT 1035
DB 211 CCAACAGAGTATAGACATGCAACGCGGCTCTCAGGCGAGGAAACAAGCTTCTGACTT 152
QY 1036 GGGGTTTACAGAGCAAGACCCCAACATGTAAGCTGGGGGGTGGGGGCTTGAATGACGGTT 1095
DB 151 GGGGTTTACAGAGCAAGACCCCAACATGTAAGCTGGGGGGTGGGGGCTTGAATGACGGTT 92
QY 1096 CTAGACTCAGGCTCAGCTCTCGGCTCTTTTCTTTTCCCTTCATTAATCCAAACCTTTT 1155

```

```

DB 91 CTAGACTCAGGCTCAGCTCTCGGCTCTTTTCTTTTCCCTTCATTAATCCAAACCTTTT 32
QY 1156 TATTCA 1161
DB 31 TATTCA 26

RESULT 10
ADP94916
ID ADP94916 standard; cDNA; 1114 BP.
XX
AC ADP94916;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human gene 11-derived ACRP30-like cDNA, SEQ ID NO:12.
XX
KW Human; ACRP30-like; adipocyte complement related protein.30-like;
KW complement factor C1q homologue; cerebellin homologue;
KW hibernation-specific protein homologue; TNF-alpha homologue;
KW tumour necrosis factor-alpha homologue; type 2 diabetes; insulin resistance;
KW drug screening; diabetes mellitus; type 2 diabetes; insulin resistance;
KW obesity; metabolic disorder; endocrine disorder; antidiabetic; anorectic;
KW gene therapy; chromosome 11q11; gene; ss.
XX
OS Homo sapiens.
XX
PN M02003031586-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002MO-US032432.
XX
PR 12-OCT-2001; 2001US-0328419P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (BLON/) BLONDEL O.
PA (RUBE/) RUBEN S M.
XX
PI Blondel O, Ruben SM;
XX
DR WPI; 2003-381707/36.
XX
PT P-PSDB; ADP94857.
XX
FT New ACRP30-like nucleic acid molecules and polypeptides, useful for
PT preventing, treating or ameliorating a medical condition, such as insulin
PT resistance or diabetes mellitus.
XX
PS Claim 1; SEQ ID NO 12; 602bp; English.
XX
CC The invention relates to novel human adipocyte complement related protein
CC 30 (ACRP30)-like cDNAs derived from 20 genes, and their encoded proteins.
CC The secretion of ACRP30 is promoted by insulin, and ACRP30 acts in a
CC similar manner to insulin, such as reducing blood glucose levels and
CC reducing levels of plasma free fatty acids. ACRP30 is homologous to
CC complement factor C1q, cerebellin, and hibernation-specific proteins HP-
CC 20, HP-25, and HP-27, and has a C-terminal globular domain similar in
CC crystal structure to C1q and tumour necrosis factor-alpha (TNF-alpha);
CC the latter of which has been implicated in insulin resistance in obesity
CC and type 2 diabetes. Additionally ACRP30 is highly expressed in
CC adipocytes during adipogenesis, and may also have immune-related
CC functions. ACRP30-like proteins may therefore be useful in the treatment
CC of insulin-related disorders such as diabetes and obesity. The invention
CC also relates to recombinant vectors and host cells comprising an ACRP30-
CC like nucleic acid; the recombinant production of ACRP30-like proteins;
CC antibodies specific for ACRP30-like proteins; use of ACRP30-like nucleic
CC acids or proteins in preventing, treating or ameliorating a medical
CC condition in an individual; methods of diagnosing a pathological
CC condition, or a susceptibility to a pathological condition; methods of
CC screening for binding partners of ACRP30-like proteins; and methods of
CC screening for modulators of ACRP30-like activity. The ACRP30-like nucleic
CC acids and proteins, and methods of the invention are useful for
CC preventing, treating or ameliorating a medical condition such as diabetes

```



CC mellitus, insulin resistance, obesity, and other metabolic and endocrine  
 CC disorders. The present sequence is related to the invention.  
 XX

Sequence 1114 BP; 243 A; 351 C; 327 G; 193 T; 0 U; 0 Other;

Query Match 93.3%; Score 1082.8; DB 10; Length 1114;  
 Best Local Similarity 99.8%; Pred. No. 1,76-244;  
 Matches 1084; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 76 GAGCTTCTTGTCTGCGAGCCCTGAGACGTCGCGGGGAGCCGCGAGAGGTAAACCAATG 135
Db 9 GAGCTTCTTGTCTGCGAGCCCTGAGACGTCGCGGGGAGCCGCGAGAGGTAAACCAATG 68

Qy 136 ATCCCTGAGGTGCTCTGAGCTGACCTCCCTGTGCTGTGACCACTGCTTGCGCC 195
Db 69 ATCCCTGAGGTGCTCTGAGCTGACCTCCCTGTGCTGTGACCACTGCTTGCGCC 128

Qy 196 TTTCCTGCGAGGACCTTCGGAAGGCTCCCTCACTGTGTGCAAGCTTGCGCC 255
Db 129 TTTCCTGCGAGGACCTTCGGAAGGCTCCCTCACTGTGTGCAAGCTTGCGCC 188

Qy 256 CAGGGCCACCCGCGCCCGCGAGCCCGAGGGCCCTCAGGAATGATGGAGCGATGGCC 315
Db 189 CAGGGCCACCCGCGCCCGCGAGCCCGAGGGCCCTCAGGAATGATGGAGCGATGGCC 248

Qy 316 TTTCCTGCGAGGACCTTCGGAAGGCTCCCTCACTGTGTGCAAGCTTGCGCC 375
Db 249 TTTCCTGCGAGGACCTTCGGAAGGCTCCCTCACTGTGTGCAAGCTTGCGCC 308

Qy 376 GGTCCACTGCGCGGACGATTAACCGGGGAAAGCCAGACCAAGGCGAAACCGGGCC 435
Db 309 GGTCCACTGCGCGGACGATTAACCGGGGAAAGCCAGACCAAGGCGAAACCGGGCC 368

Qy 436 ATTGGGGGGGCTGGCCCGCGGCGCCCAAGGGGGTCAACGGTACCCCGGGAGCATGGC 495
Db 369 ATTGGGGGGGCTGGCCCGCGGCGCCCAAGGGGGTCAACGGTACCCCGGGAGCATGGC 428

Qy 496 ACAACGAGCAAGAGGGGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 555
Db 429 ACAACGAGCAAGAGGGGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 488

Qy 556 AGCTGTGCGAGTGGCTTACCAAGTCAAGCTTCTCGGTGCGAGTGCACCAAGACTACCA 615
Db 489 AGCTGTGCGAGTGGCTTACCAAGTCAAGCTTCTCGGTGCGAGTGCACCAAGACTACCA 548

Qy 616 CGGGAGGGGCTGCCATCAAGTTTGAACAAGATTCTGATGAACGAGGGTGGCACTTAAT 675
Db 549 CGGGAGGGGCTGCCATCAAGTTTGAACAAGATTCTGATGAACGAGGGTGGCACTTAAT 608

Qy 676 GCTTCCAGCGGCAAGTCTGTGCGGGTGTGCTGGATCTTCACTTCACTTACGACATC 735
Db 609 GCTTCCAGCGGCAAGTCTGTGCGGGTGTGCTGGATCTTCACTTCACTTACGACATC 668

Qy 736 AGCTGTGCGAGGACCTTCGGAAGGCTCCCTCACTGTGTGCAAGCTTGCGCC 795
Db 669 AGCTGTGCGAGGACCTTCGGAAGGCTCCCTCACTGTGTGCAAGCTTGCGCC 728

Qy 796 ACCCTTGAATGCAACACCGGCAACGAGATGGCTCAGGCTCACCATCTTGCGCTTC 855
Db 729 ACCCTTGAATGCAACACCGGCAACGAGATGGCTCAGGCTCACCATCTTGCGCTTC 788

Qy 856 AAGCAGGAGTGAAGAGTTTGGCTGCAATCTTCACTCAAGCAAGAGCGGCTTCTAT 915
Db 789 AAGCAGGAGTGAAGAGTTTGGCTGCAATCTTCACTCAAGCAAGAGCGGCTTCTAT 848

Qy 916 GACCTTACTGAGACAGACGCTCTTAAACGGGCTTCTTAATCTATGCCAGACAGATGAC 975
Db 849 GACCTTACTGAGACAGACGCTCTTAAACGGGCTTCTTAATCTATGCCAGACAGATGAC 908

Qy 976 CCAGCAGAGATGAGATGCAAGGCGGCTCTCAAGGCAAGGAGCAAGCTTCTGAGCTT 1035
Db 909 CCAGCAGAGATGAGATGCAAGGCGGCTCTCAAGGCAAGGAGCAAGCTTCTGAGCTT 968
  
```

```

Qy 1036 GGGCTTACAGAGCAAGACCCACACTGTAGGCTGGGGGTGGGGGTGAGTAGCGGTT 1095
Db 969 GGGCTTACAGAGCAAGACCCACACTGTAGGCTGGGGGTGGGGGTGAGTAGCGGTT 1028

Qy 1096 CTAGCTCAGGCTCAGCTCTCGGCTCTTTTTCCTTCAATTAATCAACCTTTT 1155
Db 1029 CTAGCTCAGGCTCAGCTCTCGGCTCTTTTTCCTTCAATTAATCAACCTTTT 1088

Qy 1156 TATTCA 1161
Db 1089 TATTCA 1094
  
```

RESULT 11  
 ACC42468  
 ID ACC42468 standard; DNA; 993 BP.

AC42468;  
 26-AUG-2003 (first entry)  
 Human C10SF2/5 coding sequence.

Human, antiinflammatory; immunosuppressive; antipsoriatic; thyromimetic;  
 cyostatic; antimicrobial; antiaesthetic; C10SF3; C10SF; cancer;  
 Tumour Necrosis Factor; TNF; Complement C1q; inflammatory disorder;  
 immune disorder; autoimmune disorder; infection; cell adhesion; asthma;  
 psoriasis; Hashimoto's thyroiditis; inflammatory bowel disease; gene;  
 Crohn's disease; cell activity; cell behaviour; cell development; ds.

Homo sapiens.

Key Location/Qualifiers  
 CDS 1..993  
 /tag= a  
 /product= "C10SF2/5"

W02003022992-A2.  
 20-MAR-2003.

04-SEP-2002; 2002MO-US027974.  
 06-SEP-2001; 2001US-0317988P.

(SCHE ) SCHERING CORP.

Baeham BE, Forsythe IJ, Gorman DM, Mattson JD, Moshrefi M;  
 Parham CJ

WPI; 2003-313239/30.  
 P-PSDB; ABP70817.

Novel isolated human or murine C10SF3 polypeptide, and antibodies  
 specific to the polypeptides, which are useful for treating psoriasis,  
 metastasis, Hashimoto's thyroiditis, asthma, or inflammatory bowel  
 disease.

Disclosure; Page 45-46; 97pp; English.

The present invention relates to human and murine C10SF3 proteins  
 (ABP70819 and ABP70820) and their coding sequences (ACC42470 and  
 ACC42471). C10SF proteins are members of the Tumour Necrosis Factor (TNF)  
 and Complement C1q family. The C10SF proteins are useful for treating a  
 subject suffering from inflammatory disorder, immune disorder or  
 autoimmune disorder. The C10SF proteins are also useful for treating  
 infection, abnormal proliferation, cancer, metastasis, or pathological  
 cell adhesion. The C10SF proteins are also useful for treating psoriasis,  
 Hashimoto's thyroiditis, asthma, or inflammatory bowel disease, e.g.,  
 Crohn's disease. The C10SF proteins are also useful for modulating the  
 activity, behaviour, and development of cells, including monocytes, mast  
 cells, dendritic cells, macrophages, lymphocytes or neutrophils. The  
 protein encoded by the present sequence was used in a sequence alignment

XX	Sequence	with the C1QSF3 proteins	Score	Length	DB	Query Match	Best Local Similarity	Mismatches	Indels	Gaps
SO	Sequence 993 BP; 203 A; 324 C; 309 G; 157 T; 0 U; 0 Other;		85.1%;	993;	DB 8;	Query Match	99.9%;	0;	1;	0
			Pred. No. 2.4e-222;			Matches	989;	0;	1;	0
QY	1	GGAAAACATAAGCTTGGGGCCGACGCTCTGCCCCGCTGCTGCGCTGAGGAAAGCCGGGAC	60							
DB	4	GGAAAACATAAGCTTGGGGCCGACGCTCTGCCCCGCTGCTGCGCTGAGGAAAGCCGGGAC	63							
QY	61	GGGGAGCCCCCGCGAGAGCTTTGCTCCCGAGCGCCCTGACGCTGGAGGGGACGCGG	120							
DB	64	GGGAGCCCCCGCGAGAGCTTTGCTCCCGAGCGCCCTGACGCTGGAGGGGACGCGG	123							
QY	121	AGGGTAACCAACCATATATCCCTGGGGTGTCTCTGGCTGTGACCTCCCTGTGCTGTGAC	180							
DB	124	AGGGTAACCAACCATATATCCCTGGGGTGTCTCTGGCTGTGACCTCCCTGTGCTGTGAC	183							
QY	181	CCACTGCTTGGGGCTTTGCTGTGCAAGGACCTTCCGAAAGGCTCCCTCAACTGTGCTGC	240							
DB	184	CCACTGCTTGGGGCTTTGCTGTGCAAGGACCTTCCGAAAGGCTCCCTCAACTGTGCTGC	243							
QY	241	AGCGTCCCTGGGCCCCCAAGGGCCCAACCGGCCCCCAAGAGGCCCAAGGGCCCTCAAGAAAG	300							
DB	244	AGCGTCCCTGGGCCCCCAAGGGCCCAACCGGCCCCCAAGAGGCCCTCAAGAAAG	303							
QY	301	ATGGGACCAATAGGGCTTTCTTGGCAAAAGACGCGCAAGATGAGCAACAACGAGCAACGGGG	360							
DB	304	ATGGGACCAATAGGGCTTTCTTGGCAAAAGACGCGCAAGATGAGCAACAACGAGCAACGGGG	363							
QY	361	GACAGCGGAGAGAAAGGTCCACTGGCGCGGACAGGTAAACGGGGGAAAGCCAGAACCAAG	420							
DB	364	GACAGCGGAGAGAAAGGTCCACTGGCGCGGACAGGTAAACGGGGGAAAGCCAGAACCAAG	423							
QY	421	GGCAAAAGCCGGGGCCATTGGGGGGGCTGGCCCCCGTGGCCCAAGGGGGTCAACGTTACC	480							
DB	424	GGCAAAAGCCGGGGCCATTGGGGGGGCTGGCCCCCGTGGCCCAAGGGGGTCAACGTTACC	483							
QY	481	CCCGGGAAAGCATGGGCAACAAGGCAAGAAAGGGGGCCAAAGGGGAAAGGGAGCCAGGC	540							
DB	484	CCCGGGAAAGCATGGGCAACAAGGCAAGAAAGGGGGCCAAAGGGGAAAGGGAGCCAGGC	543							
QY	541	CTCCAGGCCCCCTGAGCTGTGGCAGTGGCCATACCAAGTCAGCTTTCTCGGTGGCAGTG	600							
DB	544	CTCCAGGCCCCCTGAGCTGTGGCAGTGGCCATACCAAGTCAGCTTTCTCGGTGGCAGTG	603							
QY	601	ACCAAGAGCTAACCCACGGGAGCGGCTGCCCATCAAGTTGACAAAGTTCTGATGAACGAG	660							
DB	604	ACCAAGAGCTAACCCACGGGAGCGGCTGCCCATCAAGTTGACAAAGTTCTGATGAACGAG	663							
QY	661	GGTGGCACTAACAGTCTTCCAGGGGCAAGTGGCTGGGGGGTGGCTGGGATCTACTAC	720							
DB	664	GGTGGCACTAACAGTCTTCCAGGGGCAAGTGGCTGGGGGGTGGCTGGGATCTACTAC	723							
QY	721	TTCACTTACGACATACGCTGGCCAAACAAGCACTGGCCATCGGCTGTGTGCAACAGGC	780							
DB	724	TTCACTTACGACATACGCTGGCCAAACAAGCACTGGCCATCGGCTGTGTGCAACAGGC	783							
QY	781	CAGTACCGGATCCGGACCTTTGATGCAACACCGGCAACACAGATGTGGCTCTAGGCTCC	840							
DB	784	CAGTACCGGATCCGGACCTTTGATGCAACACCGGCAACACAGATGTGGCTCTAGGCTCC	843							
QY	841	ACCAATCCGGGCTCTAAGCAGGGGTACAGAAAGTTGGCTGGACATCTTCACTCAAGCAG	900							
DB	844	ACCAATCCGGGCTCTAAGCAGGGGTACAGAAAGTTGGCTGGACATCTTCACTCAAGCAG	903							
QY	901	AACGGGCTCTTCTATGACCTTACTGACAGACAGGCTTTTACGGGCTTCTTAATCTAT	960							
DB	904	AACGGGCTCTTCTATGACCTTACTGACAGACAGGCTTTTACGGGCTTCTTAATCTAT	963							
QY	961	GCCGACCAAGATGACCCCAAGAGGTATAG	990							

```

Db      964  GCCGACCGAGTGTGACCCCAAGAGGTATAG 993
|||||
RESULT 12
AAZ45606
AAZ45606 standard; cDNA; 932 BP.
XX
XX      AAZ45606;
AC
XX
DT      06-APR-2000 (first entry)
XX
DE      cDNA encoding polypeptide ACRP30RL, a homologue of ACRP30.
XX
KW      ACRP30RL; ACRP30 homologue; 30 kDa adipose complement related protein;
KW      splice variant; ACRP30RL; inflammation; cell proliferation; cell death;
KW      immunity; energy homeostasis; cancer; obesity; diabetes; heart disease;
KW      energy metabolism; vaccine; gene therapy; ss.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      15..872
FT      FT      /*tag=a
FT      FT      /pseudo=ACRP30RL
XX
PN      WO9559618-A1.
XX
PD      25-NOV-1999.
XX
PF      18-MAY-1999; 99WO-US010822.
XX
PR      21-MAY-1998; 98US-0086562P.
PR      29-SEP-1998; 98US-00162352.
PR      20-OCT-1998; 98US-00175501.
XX
XX      (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI      Hensley P, Li X, Hu E, Smith RF, Zhu Y;
DR      WPI; 2000-116299/10.
DR      P-PSDB; MAF54321.
XX
PT      New polypeptide, useful for treating disease related to energy
PT      homeostasis including obesity.
XX
PS      Claim 2; Page 23; 36pp; English.
XX
XX      The present sequence encodes a polypeptide, designated ACRP30RL, which
XX      is a homologue of ACRP30 (a 30 kDa adipose complement related protein).
XX      ACRP30RL is a splice variant of ACRP30RL gene, encoding a protein with a
XX      68 amino acid internal insertion relative to the ACRP30RL polypeptide.
XX      The ACRP30RL polypeptide is structurally related to other proteins of
XX      the Complement C1q/Tumour necrosis factor (TNF) family. The polypeptides
XX      are believed to play a role in inflammation, cell proliferation, cell
XX      death, immunity and energy homeostasis processes. The polypeptide may be
XX      used to screen for its agonists or antagonists. Diseases or conditions
XX      arising from altered expression or activity of ACRP30RL may be diagnosed
XX      and treated with the polynucleotides and polypeptides of the invention.
XX      These diseases or conditions include cancers, inflammation, cell death,
XX      obesity, diabetes, heart disease, cell proliferation, immunity and energy
XX      metabolism and homeostasis. The polynucleotide is also useful for
XX      chromosome mapping and tissue localization. Both the ACRP30RL
XX      polypeptide and polynucleotide may be administered to patients as vaccines
XX      or as part of a gene therapy regime respectively, to treat the above
XX      diseases
SQ      Sequence 932 BP, 200 A; 300 C; 282 G; 150 T; 0 U; 0 Other;
Query March      80.1%; Score 930.4; DB 3; Length 932;
Best Local Similarity 99.9%; Pred. No. 1e-208;
Matches 931; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 119 CGAGGTAACCAACCATGATCCCTGGGTGCTCTGGCTGTGCTCCCTGCTGCTCTG 178
DB 1 CGAGGTAACCAACCATGATCCCTGGGTGCTCTGGCTGTGCTCCCTGCTGCTCTG 60
QY 179 ACCCACTGCTTTGGGCTTTGCTGCAAGGACCTTCCGAAAGGCTCCCTCAACTGCTT 238
DB 61 ACCCACTGCTTTGGGCTTTGCTGCAAGGACCTTCCGAAAGGCTCCCTCAACTGCTT 120
QY 239 GCGAGCTGCTTTGGGCTTTGCTGCAAGGACCTTCCGAAAGGCTCCCTCAACTGCTT 298
DB 121 GCGAGCTGCTTTGGGCTTTGCTGCAAGGACCTTCCGAAAGGCTCCCTCAACTGCTT 180
QY 299 TGAATGGACCAATAGGCTTTCTGTCGAAAGACGCGCAAGATGACACGACGCGACCGG 358
DB 181 TGAATGGACCAATAGGCTTTCTGTCGAAAGACGCGCAAGATGACACGACGCGACCGG 240
QY 359 GGGACAGCGGAGAGAGAGAGTTCACCTTGCCGACAGGTAACTCCGGGAAAGCAGAGCCAA 418
DB 241 GGGACAGCGGAGAGAGAGTTCACCTTGCCGACAGGTAACTCCGGGAAAGCAGAGCCAA 300
QY 419 AGGGCAAGGCGGGGCGCATTTGGGGGCTGCGCCCGTGGCGCCCAAGGGGGTCAACGGTA 478
DB 301 AGGGCAAGGCGGGGCGCATTTGGGGGCTGCGCCCGTGGCGCCCAAGGGGGTCAACGGTA 360
QY 479 CCCCAGGAGAGATGAGCAACAGGCAAGAGGGGCGCAAGAGGAGAGAGGAGGAGGAGG 538
DB 361 CCCCAGGAGAGATGAGCAACAGGCAAGAGGGGCGCAAGAGGAGAGAGGAGGAGGAGG 420
QY 539 GCCTCCAGGCGCTTGCAGCTGTGTCAGTGGCCATACCAAGTCAAGCTTTCTCGGTGGAG 598
DB 421 GCCTCCAGGCGCTTGCAGCTGTGTCAGTGGCCATACCAAGTCAAGCTTTCTCGGTGGAG 480
QY 599 TGAACCAAGAGTACCAACGGAGAGGCGTCCCATCAAGTTTGAACAAGTTCTGATGAGAG 658
DB 481 TGAACCAAGAGTACCAACGGAGAGGCGTCCCATCAAGTTTGAACAAGTTCTGATGAGAG 540
QY 659 AGGGTGGCACTCAATGACTTTCAGCGGCAAGTTCCTGCGCGGTGCTGGAGTCTACT 718
DB 541 AGGGTGGCACTCAATGACTTTCAGCGGCAAGTTCCTGCGCGGTGCTGGAGTCTACT 600
QY 719 ACTTCACTTACGATCACTGCTGGCCAAACAGACCTTGGCATTTGGTGTGCAACAG 778
DB 601 ACTTCACTTACGATCACTGCTGGCCAAACAGACCTTGGCATTTGGTGTGCAACAG 660
QY 779 GCCAGTACCGGATCCGAGCTTTGATGCAACACCGGCAACCAAGATGGGCTCAGGCT 838
DB 661 GCCAGTACCGGATCCGAGCTTTGATGCAACACCGGCAACCAAGATGGGCTCAGGCT 720
QY 839 CCACCATCTGCTGCTCAAGAGAGGATGACGAAAGTTTGGCTGCAAGTCTTCTACTCAAGC 898
DB 721 CCACCATCTGCTGCTCAAGAGAGGATGACGAAAGTTTGGCTGCAAGTCTTCTACTCAAGC 780
QY 899 AGAAGGGGCTTTCTAAGACCTTACTGAGACAGACGCTTTTACGGGGCTTCTTAATCT 958
DB 781 AGAAGGGGCTTTCTAAGACCTTACTGAGACAGACGCTTTTACGGGGCTTCTTAATCT 840
QY 959 ATGCGGACCAAGATGACCCCAAGAGAGATGACATGCAAGGAGGCTTCCAGGAGG 1018
DB 841 ATGCGGACCAAGATGACCCCAAGAGAGATGACATGCAAGGAGGCTTCCAGGAGG 900
QY 1019 AACCAAGCTTCTGACTTGGGCTTACAGAGCAA 1050
DB 901 AACCAAGCTTCTGACTTGGGCTTACAGAGCAA 932

```

```

DE Murine C1QSF2/5 coding sequence.
XX
XX Murine; antiinflammatory; immunosuppressive; antipsoriatic; thymomimetic;
XX cytostatic; antimicrobial; antiasthmatic; C1QSF3; C1QSF; cancer;
XX Tumour Necrosis Factor; TNF; Complement C1q; inflammatory disorder;
XX Immune disorder; autoimmune disorder; infection; cell adhesion; asthma;
XX Prolastis; Hashimoto's thyroiditis; inflammatory bowel disease; gene;
XX Crohn's disease; cell activity; cell behaviour; cell development; ds.
XX
OS Mus musculus.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..885 /*tag= a
XX FT /product= "C1QSF2/5"
XX
XX W02003022992-A2.
XX
XX
XX 20-MAR-2003.
XX
XX 04-SEP-2002; 2002WC-US027974.
XX
XX
XX 06-SEP-2001; 2001US-0317988P.
XX
XX (SCHE ) SCHERRING CORP.
XX
XX Basham BE, Forsythe IJ, Gorman DM, Mattson JD, Moshrefi M;
XX Parham C;
XX
XX WPI; 2003-313239/30.
XX
XX P-PSDB; ABP70818.
XX
XX
XX Novel isolated human or murine C1QSF3 polypeptide, and antibodies
XX specific to the polypeptides, which are useful for treating prolatias,
XX metastasis, Hashimoto's thyroiditis, asthma, or inflammatory bowel
XX disease.
XX
XX Disclosure; Page 47-48; 97pp; English.
XX
XX
XX The present invention relates to human and murine C1QSF3 proteins
XX (ABP70819 and ABP70820) and their coding sequences (ACC42470 and
XX ACC42471). C1QSF proteins are members of the Tumour Necrosis Factor (TNF)
XX and Complement C1q family. The C1QSF proteins are useful for treating a
XX subject suffering from inflammatory disorder, immune disorder or
XX autoimmune disorder. The C1QSF proteins are also useful for treating
XX infection, abnormal proliferation, cancer, metastasis, or pathological
XX cell adhesion. The C1QSF proteins are also useful for treating psoriasis,
XX Crohn's disease, Hashimoto's thyroiditis, asthma, or inflammatory bowel disease, e.g.,
XX Crohn's disease. The C1QSF proteins are also useful for modulating the
XX activity, behaviour, and development of cells, including monocytes, mast
XX cells, dendritic cells, macrophages, lymphocytes or neutrophils. The
XX protein encoded by the present sequence was used in a sequence alignment
XX with the C1QSF3 proteins
XX
XX
XX Sequence 885 BP; 208 A; 253 C; 262 G; 162 T; 0 U; 0 Other;
XX
XX
XX Query Match 60.6%; Score 703.6; DB 8; Length 885;
XX Best Local Similarity 88.0%; Pred. No. 2e-155;
XX Matches 766; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
XX
XX
XX 121 AGGGTAACCAACCATGATCCCTGGGTGCTCTGGCTGTGCTCCCTGCTGCTGAC 180
XX 16 AAGGTGACCAACATGATCTCTGGGTACTCTTGGCTGTGCTCCCTTCCGTGTGCTGAC 75
XX
XX 181 CCACTGCTTTGGGCTTTGCTGCAAGGACCTTCCGAAAGGCTCCCTCAACTGCTTGC 240
XX 76 CCAATGCTTGGTGTGCTTGTCTGCAAGGACCTTCCGAAAGGAGGCTTCAACTGATTTGC 135
XX
XX 241 AGCTGCTGGGCGCCCAAGGCGCAACCGGCGCCCAAGAGAGCCCAAGGAGGCTTCAAGATG 300
XX 136 AGCTGCTGTGTGCTCCCAAGGCGCAACCTGCGCTTCCAGAGACCAAGATCTCAAGAGTG 195
XX
XX 301 ATGGGACGAATGGGCTTTCTGCGAAAGACGGCCCAAGATGAGACGACGCGGACCGGGGG 360

```

Db 196 GTGGGAAGATGGTTTCCCTGGGAAGACCGCCCAAGATGGCCAGAGACCGGGGG 255  
 Qy 361 GACAGCGGAGAGAAAGTTCACCTGGCGGACAGGTAAACCGGGGAAACCGAGCAAAAG 420  
 Db 256 GACAGTGAAGAAAGATCCACCTGGCAGACAGGCAACCGTGAAACCAAGCAAAAG 315  
 Qy 421 GGCAAGCGGGGCGCATTTGGGCGGGCTGGCCCTGGCCCGCAAGGGGTCAACGGTACC 480  
 Db 316 GGCAAGCGGGGCGCATTTGGGCGGGCTGGCCCTGGCCCGCAAGGGGTCAACGGTACC 375  
 Qy 481 CCCGGGAAGCATGACACACGAGGCAAGAGGGGCCCAAGGGGCAAGAAAGGAGCCAGGC 540  
 Db 376 CCCGGGAAGCATGACACACGAGGCAAGAGGGGCCCAAGGGGCAAGAAAGGAGCCAGGC 435  
 Qy 541 CTCCTCAGGCCCCCTGACGCTGGAGAGTGGCCATACCAAGTCACTTTCTCGTGGCAGTGG 600  
 Db 436 CTCCTCAGGCCCCCTGACGCTGGAGAGTGGCCATACCAAGTCACTTTCTCGTGGCAGTGG 495  
 Qy 601 ACCAAGAGCTACCCACGCGGAGCGCTGCCCATCAAGTTTGACAAAGATTCTGATGAACGAG 660  
 Db 496 ACCAAGAGCTACCCACGCGGAGAGACTGGCTTCAAGATTGACAAAGATTCTGATGAACGAG 555  
 Qy 661 GGTGGCCACTCAATGCTTCCAGCGGCAAGTTGCTGGCGGCTGGCTGGATCTACTAC 720  
 Db 556 GGTGGCCACTCAATGCTTCCAGCGGCAAGTTGCTGGCGGCTGGCTGGATCTACTAC 615  
 Qy 721 TTCACTTACGACATCAGCGCTGGCCCAACAGCACTGGCCATCGCGCTGGTGGCAACGAGC 780  
 Db 616 TTCACTTACGACATCAGCGCTGGCCCAACAGCACTGGCCATCGCGCTGGTGGCAACGAGC 675  
 Qy 781 CAGTACCGCATCCCGACCTTTGATGATCCCAACACCGGCAACACAGATGTGGCTCAGGCTCC 840  
 Db 676 CAGTACCGCATCCCGACCTTTGATGATCCCAACACCGGCAACACAGATGTGGCTCAGGCTCC 735  
 Qy 841 ACCATCTGAGCTCTCAAGCAGAGGTGACAGAGTTTGGCTGACAGATCTTCACTACAGAGAG 900  
 Db 736 ACCATCTGAGCTCTCAAGCAGAGGTGATAGATCTGCTGACAGATCTTCACTACAGAGAG 795  
 Qy 901 AAGCGGCTCTTCTATGACCTTTACTGACAGACAGCTCTTTACGCGGCTCTTCAATCTAT 960  
 Db 796 AATGGCTCTTCTATGACCTTTACTGACAGACAGCTCTTTACGCGGCTCTTCAATCTAT 855  
 Qy 961 GCCGACCGAGATGACCCCAACGAGGTATAG 990  
 Db 856 GCTGACCAAGAGACCCCAACGAGGTATAG 885

## RESULT 14

AA261744  
ID AA261744 standard; cDNA; 1107 BP.

AC AA261744;

DT 27-MAR-2000 (first entry)

DE cDNA encoding rat skin cell secreted protein, SEQ ID NO:217.

XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;  
 XX embryonic skin cell; keratinocyte stem cell; transit amplifying cell;  
 XX secreted; transmembrane; inflammation; cancer; neurological disease;  
 XX angiogenesis; tumour vascularisation; growth disorder;  
 XX developmental disorder; skin wound; hair follicle disorder;  
 XX anti-inflammatory; cyostatic; neuroprotective; vulnery; ss.

XX Rattus sp.

XX MO9955865-A1.

XX 04-NOV-1999.

XX 29-APR-1999; 99WO-NZ000051.

XX

PR 29-APR-1998; 98US-00069726.  
 PR 09-NOV-1998; 98US-00188930.  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;  
 DR WPI, 2000-072177/06.  
 DR P-P8DB; AAY76039.  
 PT Novel polynucleotides useful for the treatment of various conditions  
 PT including wounds and cancer.  
 PS Claim 1; Page 142-143; 235pp; English.  
 XX The invention relates to novel nucleic acid sequences derived from rat  
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,  
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying  
 CC cells. Polypeptides of the invention may be used to treat inflammation,  
 CC cancer and neurological diseases. The proteins may be used to stimulate  
 CC the growth and motility of keratinocytes, to inhibit the growth of cancer  
 CC cells, to modulate angiogenesis and tumour vascularisation, to modulate  
 CC skin inflammation, to modulate epithelial cell growth and to inhibit  
 CC binding of HIV-1 to leukocytes. The invention may also be used to treat  
 CC growth and developmental defects, skin wounds and hair follicle  
 CC disorders. Sequences AA261606-261832 represent cDNA sequences derived  
 CC from several mouse, rat or human skin cell types. Sequences AA261606-  
 CC 261649, AA261725-261765, AA261802-261811 and AA261826 encode proteins  
 CC with an N-terminal signal sequence, indicating that the proteins are  
 CC secreted. Sequences AA261650-261668, AA261766-261817 and  
 CC AA261827-261829 encode proteins with one or more putative transmembrane  
 CC domains  
 SQ Sequence 1107 BP; 273 A; 298 C; 328 G; 208 T; 0 U; 0 Other;  
 Query Match 60.5%; Score 702.2; DB 3; Length 1107;  
 Best Local Similarity 84.7%; Pred. No. 4,56-155;  
 Matches 800; Conservative 0; Mismatches 143; Indels 2; Gaps 1;  
 Qy 121 AGGGTAACCAATGATCCCTGGGATGCTCTGAGCTGTGCTCCCTGCTGCTGAC 180  
 Db 162 AAGGTGACCAACATGATCTCTGGATGCTCTGAGCTGTGCTCCCTGCTGCTGAC 221  
 Qy 181 CCACTGCTGGCGGCTTTGCTGCGAGGAGCTTCCGGAAGGCTCCCTCACTGCTGCG 240  
 Db 222 CCAATGCTTGGTGTCTTGTGCTGCGAGGAGCTTCCGGAAGGCTCTCACTGCTGCG 281  
 Qy 241 AGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 Db 282 AGTGTGCTGTGCTCCCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 341  
 Qy 301 ATGGGACGATGAGCTTCTCTGCGAAGACGCGCAAGATGACACAGCGGAGCTGGGGG 360  
 Db 342 GTGGGAAGAAATGGGTTTCTGTAGAGATGGCCAAAGCGGCGGCGGCGGCGGCGGCGG 401  
 Qy 361 GACAGCGGAGAGAAAGTTCACCTGGCGGACAGTAAACGGGGAAAGCCAGAACCAAG 420  
 Db 402 GACAGTGAAGAAAGTTCACCTGGCGGACAGTAAACGGGGAAAGCCAGAACCAAG 461  
 Qy 421 GGCAAGCGGGGCGCATTTGGGCGGGCTGGCCCGCGTGGCCCGCAAGGGGCTCAACGGTACC 480  
 Db 462 GGCAAGCGGGGCGCATTTGGGCGGGCTGGCCCGCGTGGCCCGCAAGGGGCTCAACGGTACC 521  
 Qy 481 CCCGGGAAGCATGACACACGAGGCAAGAGGGGCCCAAGGGGCAAGAAAGGAGCCAGGC 540  
 Db 522 CCCGGGAAGCATGATATCCGGGCAAGAGGAGCTTAAGGGCAAGAAAGGAGGAACTGGG 581  
 Qy 541 CTCCTCAGGCCCCCTGACGCTGGAGAGTGGCCATACCAAGTCACTTTCTCGTGGCAGTGG 600  
 Db 582 CTCCTCAGGCCCCCTGACGCTGGAGAGTGGCCATACCAAGTCACTTTCTCGTGGCAGTGG 641  
 Qy 601 ACCAAGAGCTACCCACGCGGAGCGGCTGCCATCAAGTTTGACAAAGATTCTGATGAACGAG 660



Qy 1019 AACAGCTTCTGCACTTGGGCTTACAGACAGACCCCACTG 1063  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1062 ACTAAGATTCCGCAAGGGTGTCTGATGAGAGAGATCTCTGAAC TG 1106

Search completed: December 25, 2004, 04:01:21  
Job time : 663 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2004, 02:34:02 : Search time 4236 Seconds  
(without alignments)  
9987.377 Million cell updates/sec

Title: US-10-621-787-1  
Perfect score: 1161  
Sequence: 1 ggaataactagcctgggggc.....aatccaactttattca 1161

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues  
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_est8: \*  
9: gb\_est9: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	779.4	67.1	863	5	BO881020 AGENCOURT
2	716.2	61.7	1196	3	AK007683 Mus muscu
3	645.2	55.6	671	5	BQ006349 UI-H-E11-
4	630.4	54.3	654	4	BM675363 UI-E-E10-
5	607.8	52.4	806	4	BI109819 602901018
6	597	51.4	634	4	BM714600 UI-E-E10-
7	588.2	50.7	708	4	BI489428 603020916
8	547	47.1	714	6	CF108213 Shultzomi
9	535.4	46.1	544	1	AI916779 wd26e01.x
10	523.8	45.1	535	2	AM304481 xv22h08.x
11	500.6	43.1	659	4	BI822472 603038061
12	493	42.5	557	1	AA732948 z978d10.8
13	466	40.1	904	4	BI488418 603020916
14	458.8	39.5	526	4	BG381862 296753 MA
15	452.4	39.0	455	1	AI288625 q188b05.x
16	438.6	37.8	496	2	BE232317 137306 MA
17	432	37.2	559	1	AV616838 AV616838
18	413.8	35.6	639	5	BO207443 UI-R-DY1-
19	413.4	35.6	556	2	AM965223 EST377296
20	398.6	34.3	444	1	AA443244 aa14c05.x
21	397.6	34.2	696	6	BQ210972 UI-R-DY1-
22	392.8	33.8	715	7	CK776809 968175 MA
23	388.6	33.5	670	6	CB324225 UI-R-DY0-
24	387.4	33.4	826	7	CK794745 AGENCOURT

25	383.2	33.0	663	6	CB323171	CB323171 UI-R-DY0-
26	373.2	32.1	425	1	AA194106	AA194106 ZR18G03.8
27	369.6	31.8	528	4	BG670534	BG670534 DRNBCC10
28	356	30.7	449	8	A2805724	A2805724 ZM0067H16
29	338.4	29.1	409	7	W79527	W79527 z481604.81
30	333	28.7	351	1	AI038644	AI038644 OX39d11.8
31	326.8	28.1	344	1	AA443206	AA443206 aa14c06.8
32	326.8	28.1	564	2	BE851196	BE851196 uw93d01.Y
33	311.8	26.9	374	7	W41045	W41045 mc35b09.r1
34	307.4	26.5	870	9	AY420023	AY420023 Mus muscu
35	305.8	26.3	1436	3	AK078818	AK078818 Mus muscu
36	304.4	26.2	463	6	CB735431	CB735431 AMGNNUC:N
37	295.4	25.4	870	9	AY420021	AY420021 Homo sapi
38	286.8	24.7	502	4	BI303579	BI303579 UI-R-DY0-
39	284	24.5	451	6	CB787179	CB787179 AMGNNUC:N
40	274.6	23.7	583	4	BI696475	BI696475 603347416
41	272.4	23.5	483	5	BX527335	BX527335 BX527335
42	269.6	23.2	870	9	AY420022	AY420022 Pan trogl
43	269.4	23.2	293	2	BF511014	BF511014 UI-H-B14-
44	264.8	22.8	714	5	BU127923	BU127923 603114908
45	263.2	22.7	734	5	BU323601	BU323601 603492824

#### ALIGNMENTS

RESULT 1  
BO881020  
LOCUS  
DEFINITION  
AGENCOURT 8124216 Lupski dorsal root ganglion Homo sapiens CDNA  
clone IMAGE:6178402 5', mRNA sequence.  
BO881020  
BO881020.1 GI:22273028  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Unpublished (1999)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LHAM13557 row: 0 column: 11  
High quality sequence stop: 722.  
Location/Qualifiers

#### FEATURES

1. 863  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6178402"  
/sex="male"  
/tissue\_type="dorsal root ganglia"  
/dev\_stage="adult", 36 yr  
/lab\_host="DH10B"  
/clone\_lib="Lupski dorsal root ganglion"  
/note="Vector: PCMV-SPORT6 (Life Technologies); Site 1:  
NotI; Site 2: SalI; CDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCACCCACGCTCCG-3' and  
5'-GACTAGTCTTAGTCGACGCGCCCTT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life

ORIGIN Technologies."

Query Match 67.1%; Score 779.4; DB 5; Length 863;

Best Local Similarity 97.3%; Pred. No. 2.4e-177;

Matches 825; Conservative 0; Mismatches 17; Indels 6; Gaps 3;

24 GCTCTGCGCGCGCTGCTGCGCGCTGAGAAAGCGCGGAGCGCGCGCGCGAGGCTTCT 83  
 1 GCTTGGCGCGCGCTGCTGCGCGCTGAGAAAGCGCGGAGCGCGCGCGCGAGGCTTCT 60  
 84 TTGCTCGGAGCGCGCTGAGAGTGGCGGAGCGCGGAGGTTAAACAACATGATCCCTG 143  
 61 TTGCTCGGAGCGCGCTGAGAGTGGCGGAGCGCGGAGGTTAAACAACATGATCCCTG 120  
 144 GGTGCTCTGCGCGCTGCTGCGCGCTGCTGCTGAGAACCTGCTTGGCGCTTGGCTCG 203  
 121 GGTGCTCTGCGCGCTGCTGCGCGCTGCTGCTGAGAACCTGCTTGGCGCTTGGCTCG 180  
 204 CAGGAGACTTCGCGGAAAGGCTCCCTCAACTGCTGAGCTGCTGAGCGCGCGCGCGCC 263  
 181 CAGGAGACTTCGCGGAAAGGCTCCCTCAACTGCTGAGCTGCTGAGCGCGCGCGCGCC 240  
 264 ACCCGCGCGCGCGCGGAGCGCGCGCGCTCTAGAAATGATGAGAGGAGCTTCTCG 323  
 241 ACCCGCGCGCGCGCGGAGCGCGCGCGCTCTAGAAATGATGAGAGGAGCTTCTCG 300  
 324 CAAAGAGCGCGCAATAGACACGAGCGCGAGCGCGCGCGCGCGAGAGAGGTTCAACC 383  
 301 CAAAGAGCGCGCAATAGACACGAGCGCGAGCGCGCGCGCGCGAGAGAGGTTCAACC 360  
 384 TGGCGCGGACAGGTACCGCGGAGAGCGAGGACCAAGAGGCGCGCGCGCGCATTTGGAGC 443  
 361 TGGCGCGGACAGGTACCGCGGAGAGCGAGGACCAAGAGGCGCGCGCGCATTTGGAGC 420  
 444 GGTGCGCGCGCGCGCGGAGCGCGCGGTCAACGCTAACCGCGCGGAGAGGATGACACGAG 503  
 421 GGTGCGCGCGCGCGCGGAGCGCGGTCAACGCTAACCGCGCGGAGAGGATGACACGAG 480  
 504 CAAAGAGCGCGCGCAAGAGGAGGAGAGGAGCGCGCGCTCCAGAGCGCGCGCGCGCTGAG 563  
 481 CAAAGAGCGCGCGCAAGAGGAGGAGGAGGAGCGCGCGCTCCAGAGCGCGCGCGCTGAG 540  
 564 CAGTGGCGCATACCAAGTCAAGCTTCTCGTGGAGGAGTACCAAGAGTACCAAGAGGAGCG 623  
 541 CAGTGGCGCATACCAAGTCAAGCTTCTCGTGGAGGAGTACCAAGAGTACCAAGAGGAGCG 600  
 624 GCTGCGCATACCAAGTTCAGAAATTCGATGAAACGAGGAGTGGCACTACCAATGCTTCAG 683  
 601 GCTGCGCATACCAAGTTCAGAAATTCGATGAAACGAGGAGTGGCACTACCAATGCTTCAG 660  
 684 CGGGAATTCGTCGCGCGCGCGCTGGAGTCTACTCTTCACTTCAAGCAATCAGGCTGGC 743  
 661 CGGGAATTCGTCGCGCGCGCGCTGGAGTCTACTCTTCACTTCAAGCAATCAGGCTGGC 720  
 744 CAAAGAGCGCGCATACCGCGCTGCTGAGCAAGAGCGCGCATGACCGCATCC-GAGCTTTG 802  
 721 CAAAGAGCGCGCATACCGCGCTGCTGAGCAAGAGCGCGCATGACCGCATCCGGAGCTTTG 780  
 803 ATGCGCAACACCGG---CAACCAAGATGAGGCTTCAGGCTTCACCAT-CTGAGCTTCGAA 857  
 781 ATGCGCAACACCGGCAACCAAGATGAGGCTTCAGGCTTCACCATCCCTGCTTCGAA 840  
 858 GCGAGGCG 865  
 841 GAAAGGCG 848

RESULT 2  
 AK007683 1196 bp mRNA linear HTC 03-APR-2004  
 LOCUS AK007683  
 DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length  
 enriched library, clone:1810033K05 product:COMPLEMENT-C10 TUMOR

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE

AUTHORS

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

NECROSIS FACTOR-RELATED PROTEIN (COMPLEMENT-C10 TUMOR NECROSIS  
 FACTOR-RELATED PROTEIN 2) homolog [Homo sapiens], full insert  
 sequence.

AK007683

AK007683.1 GI:12841383

HTC; CAP trapper.

Mus musculus

Mus musculus (house mouse)

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,

Komuro, H., Akiyama, J., Niehi, K., Kikunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Harada, M., Nishino, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format

sequencing pipeline with 384 multiplexed capillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the

PANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

The PANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 1196)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,

Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,

Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,

Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komuro, H., Kouda, M.,

Koyama, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,

Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,

Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,

Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,

Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,

Tejima, Y., Toya, T., Yamamuro, T., Yasunishi, A., Yoshida, K.,

Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Direct SubMISSION

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,

URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

Please visit our web site (http://genome.gsc.riken.jp/) for further

details.

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken





ORIGIN TAG\_SEQ=ACCTA"

Query Match 54.3%; Score 630.4; DB 4; Length 654;  
Best Local Similarity 99.8%; Pred. No. 2.2e-141;  
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

530 GGGAGCAGGCGCTCCAGAGCCCTGACGCTGTGACAGTGCATACCAAGTCACTTTCT 589  
530 GGGAGCAGGCGCTCCAGAGCCCTGACGCTGTGACAGTGCATACCAAGTCACTTTCT 589  
653 GGGAGCAGGCGCTCCAGAGCCCTGACGCTGTGACAGTGCATACCAAGTCACTTTCT 594  
653 GGGAGCAGGCGCTCCAGAGCCCTGACGCTGTGACAGTGCATACCAAGTCACTTTCT 594  
590 CGGTGAGTGAACCAAGAGTACCAAGCGGAGCGGCTGCCATCAAGTTTGAAGAATTC 649  
590 CGGTGAGTGAACCAAGAGTACCAAGCGGAGCGGCTGCCATCAAGTTTGAAGAATTC 649  
593 CGGTGAGTGAACCAAGAGTACCAAGCGGAGCGGCTGCCATCAAGTTTGAAGAATTC 534  
593 CGGTGAGTGAACCAAGAGTACCAAGCGGAGCGGCTGCCATCAAGTTTGAAGAATTC 534  
650 TGATGAACGAGGCTGCGCACTACATGCTTCAGCGGCAAGTTGCTGCGGCTGCTG 709  
650 TGATGAACGAGGCTGCGCACTACATGCTTCAGCGGCAAGTTGCTGCGGCTGCTG 709  
533 TGATGAACGAGGCTGCGCACTACATGCTTCAGCGGCAAGTTGCTGCGGCTGCTG 474  
533 TGATGAACGAGGCTGCGCACTACATGCTTCAGCGGCAAGTTGCTGCGGCTGCTG 474  
710 GGAATCTACTTCACTAGACATCAAGCTGCGGCAACAGCACTGCGCATCGGCTG 769  
710 GGAATCTACTTCACTAGACATCAAGCTGCGGCAACAGCACTGCGCATCGGCTG 769  
473 GGAATCTACTTCACTAGACATCAAGCTGCGGCAACAGCACTGCGCATCGGCTG 414  
473 GGAATCTACTTCACTAGACATCAAGCTGCGGCAACAGCACTGCGCATCGGCTG 414  
770 TGACAAACGCGCAAGTACCGCATCCGACCTTTGATGCAACACCGCAACCAAGTGG 829  
770 TGACAAACGCGCAAGTACCGCATCCGACCTTTGATGCAACACCGCAACCAAGTGG 829  
413 TGACAAACGCGCAAGTACCGCATCCGACCTTTGATGCAACACCGCAACCAAGTGG 354  
413 TGACAAACGCGCAAGTACCGCATCCGACCTTTGATGCAACACCGCAACCAAGTGG 354  
830 CTTGAGCTCCACCATCTGCTGCTCAAGCAGGCTGACGAAGTTTGCTGCAATCTTCT 889  
830 CTTGAGCTCCACCATCTGCTGCTCAAGCAGGCTGACGAAGTTTGCTGCAATCTTCT 889  
353 CTTGAGCTCCACCATCTGCTGCTCAAGCAGGCTGACGAAGTTTGCTGCAATCTTCT 294  
353 CTTGAGCTCCACCATCTGCTGCTCAAGCAGGCTGACGAAGTTTGCTGCAATCTTCT 294  
890 ACTCAGAGCAAGAGCGGCTCTTCTGACCTTCTGACAGAGCTCTTTACGGGCT 949  
890 ACTCAGAGCAAGAGCGGCTCTTCTGACCTTCTGACAGAGCTCTTTACGGGCT 949  
293 ACTCAGAGCAAGAGCGGCTCTTCTGACCTTCTGACAGAGCTCTTTACGGGCT 234  
293 ACTCAGAGCAAGAGCGGCTCTTCTGACCTTCTGACAGAGCTCTTTACGGGCT 234  
950 TCTTAATCTATGCGACAGAGTACCCCAAGAGTATGACATGCAACGCGGCTCTC 1009  
950 TCTTAATCTATGCGACAGAGTACCCCAAGAGTATGACATGCAACGCGGCTCTC 1009  
233 TCTTAATCTATGCGACAGAGTACCCCAAGAGTATGACATGCAACGCGGCTCTC 174  
233 TCTTAATCTATGCGACAGAGTACCCCAAGAGTATGACATGCAACGCGGCTCTC 174  
1010 CAGGCAAGGAAACAAAGCTTGTGACCTTGTGACGAGCAAGCAAGCAAGCTGAGCT 1069  
1010 CAGGCAAGGAAACAAAGCTTGTGACCTTGTGACGAGCAAGCAAGCAAGCTGAGCT 1069  
173 CAGGCAAGGAAACAAAGCTTGTGACCTTGTGACGAGCAAGCAAGCAAGCTGAGCT 114  
173 CAGGCAAGGAAACAAAGCTTGTGACCTTGTGACGAGCAAGCAAGCAAGCTGAGCT 114  
1070 GGGGCTGGGGGCTGAGTGAAGCGGCTTCAAGCTCAGGCTCAAGCTCCTCTTTT 1129  
1070 GGGGCTGGGGGCTGAGTGAAGCGGCTTCAAGCTCAGGCTCAAGCTCCTCTTTT 1129  
113 GGGGCTGGGGGCTGAGTGAAGCGGCTTCAAGCTCAGGCTCAAGCTCCTCTTTT 54  
113 GGGGCTGGGGGCTGAGTGAAGCGGCTTCAAGCTCAGGCTCAAGCTCCTCTTTT 54  
Qy 1130 TCCCTTCATTAATCCAAACCTTTTATTCA 1161  
Qy 1130 TCCCTTCATTAATCCAAACCTTTTATTCA 1161  
Db 53 TCCCTTCATTAATCCAAACCTTTTATTCA 22  
Db 53 TCCCTTCATTAATCCAAACCTTTTATTCA 22

RESULT 5  
B1109819 806 bp mRNA linear EST 26-JUN-2001  
LOCUS 602901018P1 NCI\_GCAP\_Mam5 Mus musculus cDNA clone IMAGE:5030717 5',  
DEFINITION mRNA sequence.  
ACCESSION B1109819  
VERSION B1109819.1 GI:14560720  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 806)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strauberg, Ph.D.  
Email: cga@nci.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LMNL at:  
http://image.lnl.gov  
Plate: LMNL1086 row: C column: 06  
High quality sequence stop: 806.  
Location/Qualifiers

## FEATURES

source

1. 806  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="mix FVB/N, C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5030717"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/clone\_id="NCI\_GCAP Mam5"  
/note="Organ: mammary; Vector: pCMV-Sport6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

## ORIGIN

Query Match 52.4%; Score 607.8; DB 4; Length 806;  
Best Local Similarity 88.0%; Pred. No. 6.7e-136;  
Matches 696; Conservative 0; Mismatches 92; Indels 3; Gaps 3;

174 TGCTGACCACTGCTTGGCGCTTGTGCGAGGAACTTCGGAAGGCTCCCTCAACT 233  
174 TGCTGACCACTGCTTGGCGCTTGTGCGAGGAACTTCGGAAGGCTCCCTCAACT 233  
1 TGCTGACCACTGCTTGGCGCTTGTGCGAGGAACTTCGGAAGGCTCCCTCAACT 60  
1 TGCTGACCACTGCTTGGCGCTTGTGCGAGGAACTTCGGAAGGCTCCCTCAACT 60  
234 GGTCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 293  
234 GGTCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 293  
61 AGTTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
61 AGTTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
294 AGAATGATGGAAGCAATGAGGCTTCTGCGCAAGAGCGGCAAGATGACAGAGCGA 353  
294 AGAATGATGGAAGCAATGAGGCTTCTGCGCAAGAGCGGCAAGATGACAGAGCGA 353  
121 AGAATGATGGAAGCAATGAGGCTTCTGCGCAAGAGCGGCAAGATGACAGAGCGA 180  
121 AGAATGATGGAAGCAATGAGGCTTCTGCGCAAGAGCGGCAAGATGACAGAGCGA 180  
354 CCGGGGGGACAGCGGAGAGGAGGTCACCTGCGGACAGAGTAAACCGGGGAAACCAAG 413  
354 CCGGGGGGACAGCGGAGAGGAGGTCACCTGCGGACAGAGTAAACCGGGGAAACCAAG 413  
181 CCGGGGGGACAGCGGAGAGGAGGTCACCTGCGGACAGAGTAAACCGGGGAAACCAAG 240  
181 CCGGGGGGACAGCGGAGAGGAGGTCACCTGCGGACAGAGTAAACCGGGGAAACCAAG 240  
414 ACCAAAGGCAAGACCGGGGCTATTGGGCGGCTGCGGCTGCGGCTGCGGCTGCGG 473  
414 ACCAAAGGCAAGACCGGGGCTATTGGGCGGCTGCGGCTGCGGCTGCGGCTGCGG 473  
241 ACCAAAGGCAAGACCGGGGCTATTGGGCGGCTGCGGCTGCGGCTGCGGCTGCGG 300  
241 ACCAAAGGCAAGACCGGGGCTATTGGGCGGCTGCGGCTGCGGCTGCGGCTGCGG 300  
474 CCGTACCCCGGGGAGAGATGACACACAGGCAAGAGGCGGCAAGAGGAGGAGGAG 533  
474 CCGTACCCCGGGGAGAGATGACACACAGGCAAGAGGCGGCAAGAGGAGGAGGAG 533  
301 TGATACCCCGGGGAGAGATGACACACAGGCAAGAGGCGGCAAGAGGAGGAGGAG 360  
301 TGATACCCCGGGGAGAGATGACACACAGGCAAGAGGCGGCAAGAGGAGGAGGAG 360  
534 GCCAGGCTCCAGAGCCCTGACGCTGTGACAGTGCATACCAAGTCAAGCTTCTCGT 593  
534 GCCAGGCTCCAGAGCCCTGACGCTGTGACAGTGCATACCAAGTCAAGCTTCTCGT 593  
361 GCTGAGCTCCAGAGCCCTGACGCTGTGACAGTGCATACCAAGTCAAGCTTCTCGT 420  
361 GCTGAGCTCCAGAGCCCTGACGCTGTGACAGTGCATACCAAGTCAAGCTTCTCGT 420  
594 GGCAGTGAACCAAGCTACCAAGCGGAGCGGCTGCCATCAAGTTTGAAGAATTCGAT 653  
594 GGCAGTGAACCAAGCTACCAAGCGGAGCGGCTGCCATCAAGTTTGAAGAATTCGAT 653  
421 GGCAGTGAACCAAGCTACCAAGCGGAGCGGCTGCCATCAAGTTTGAAGAATTCGAT 480  
421 GGCAGTGAACCAAGCTACCAAGCGGAGCGGCTGCCATCAAGTTTGAAGAATTCGAT 480  
654 GAAAGAGGCTGAGCACTACCAATGCTTCAGAGGCAAGTTCGCTGCGGCTGAGGAT 713  
654 GAAAGAGGCTGAGCACTACCAATGCTTCAGAGGCAAGTTCGCTGCGGCTGAGGAT 713  
481 GAAAGAGGCTGAGCACTACCAATGCTTCAGAGGCAAGTTCGCTGCGGCTGAGGAT 540  
481 GAAAGAGGCTGAGCACTACCAATGCTTCAGAGGCAAGTTCGCTGCGGCTGAGGAT 540  
714 CTACTACTTCACTACAGACATCAAGCTGCGCAACAGCACTGCGCATCGGCTGAGCA 773  
714 CTACTACTTCACTACAGACATCAAGCTGCGCAACAGCACTGCGCATCGGCTGAGCA 773  
541 CTACTACTTCACTACAGACATCAAGCTGCGCAACAGCACTGCGCATCGGCTGAGCA 600  
541 CTACTACTTCACTACAGACATCAAGCTGCGCAACAGCACTGCGCATCGGCTGAGCA 600  
774 CAAGGCAAGTACCGGATCCGAGCTTGTGATGCCAACAAGGCAACCAAGTATGAGCTC 833  
774 CAAGGCAAGTACCGGATCCGAGCTTGTGATGCCAACAAGGCAACCAAGTATGAGCTC 833  
601 CAAGGCAAGTACCGGATCCGAGCTTGTGATGCCAACAAGGCAACCAAGTATGAGCTC 660  
601 CAAGGCAAGTACCGGATCCGAGCTTGTGATGCCAACAAGGCAACCAAGTATGAGCTC 660  
834 AGGCTCACCATCTGAGCTCAAC-GCAGGAGTGAAGAAAGTTTGAGCTGCAAGTCTTCACT 892  
834 AGGCTCACCATCTGAGCTCAAC-GCAGGAGTGAAGAAAGTTTGAGCTGCAAGTCTTCACT 892

Db	661	GGGCTCCACCATCTAGCTCTCAGAGCAGGGTATGAAAGTCTGGCTGCAGATCTTTACT	720
Qy	893	CAGAGCAGAAACGGGCTCTTCTATGACCTTATCTGACAGACAGCCTCTT-TACGGGCTTC	951
Db	721	CAGAGCAGAAATGGGCTCTTTACGA-CCTTACTGGACGACAGCCTGTTCACACGGCTTC	779
Qy	952	CTATCTATGC	962
Db	780	CTCATCTACGC	790
RESULT 6			
BM714600			
LOCUS			
DEFINITION	BM714600	634 bp	mRNA
ACCESSION	UI-E-EJ0-ahr-p-10-0-UI.r1	UI-E-EJ0	Homo sapiens
VERSION	BM714600		CDNA clone
KEYWORDS	BM714600.1	GI:19027858	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Mukayyotai, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	1 (bases 1 to 634)		
JOURNAL	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
MEDLINE	Normalization and Subtraction: two approaches to facilitate gene		
PubMed	discovery		
COMMENT	Genome Res. 6 (9), 791-806 (1996)		
	9704447		
	8889548		
	Contact: Soares, MB		
	Coordinated Laboratory for Computational Genomics		
	University of Iowa		
	375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA		
	Tel: 319 335 8250		
	Fax: 319 335 9555		
	Email: bento-soares@uiowa.edu		
	Tissue Procurement: Dr. Gregg Hageman		
	CDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
	CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
	Clone Distribution: Researchers may obtain clones from Research		
	Genetics (www.resgen.com).		
	Seq primer: M13 Reverse.		
FEATURES			
source	Location/Qualifiers		
	1..634		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="UI-E-EJ0-ahr-p-10-0-UI"		
	/tissue type="fetal eyes, lens, eye anterior segment,		
	optic nerve, retina, Retina Foveal and Macular, RPE and		
	Choroid"		
	/dev_stage="fetal and adult"		
	/lab_host="DHI0B (Life Technologies) (T1 phage resistant)"		
	/clone_id="UI-E-EJ0"		
	/note="Organ: eye; Vector: pT773-Pac (Pharmacia) with a		
	modified polylinker; Site 1: Ecor I; Site 2: Not I;		
	UI-E-EJ0 is a subcloned CDNA library constructed		
	according to Bonaldo, Lennon and Soares, Genome Research,		
	6:791-806, 1996. First strand CDNA synthesis was primed		
	with an oligo-dT primer containing a Not I site. Double		
	stranded CDNA was ligated to an Ecor I adaptor, digested		
	with Not I, and cloned directionally into pT773-Pac		
	vector. The oligonucleotide used to prime the synthesis of		
	first-strand CDNA contains a library tag sequence that is		
	located between the Not I site and the (dT)18 tail. The		
	sequence tags for this library are: fetal eyes,		
	AGATTCAGAG; lens, CGATTCAGCA; eye anterior segment,		
	AATCCCGCAT; optic nerve, CCATTCAGG; retina, CCGCG;		
	Foveal and Macular, GTCC; RPE and Choroid, ACTCA. This		
	library was created for the program, Gene Discovery in the		

ORIGIN	Visual System, supported by National Eye Institute (NEI).
Query Match	51.4%; Score 597; DB 4; Length 634;
Best Local Similarity	99.0%; Pred. No. 2,6e-133;
Matches 611; Conservative	0; Mismatches 5; Indels 1; Gaps 1;
OY	2 GAAAACTATGCTTGGGGCCGACGCTCTGGCCCGGTGGCTGGAGAGAAACCGGGACG 61
DB	1 GAAAACTATGCTTGGGGCCGACGCTCTGGCCCGGTGGCTGGAGAGAAACCGGGACG 60
OY	62 CGAGAGCCCGCCGAGAGCTTCTTGTCTCCGGAAGCCCTCTGGAAGCTGGACAGCCGGA 121
DB	61 CGAGAGCCCGCCGAGAGCTTCTTGTCTCCGGAAGCCCTCTGGAAGCTGGACAGCCGGA 120
OY	122 GGGTAACCAACCATGATCCCTGGGTGCTCTGGGCTTGGCCCTCTGTGTCTGAC 181
DB	121 GGGTAACCAACCATGATCCCTGGGTGCTCTGGGCTTGGCCCTCTGTGTCTGAC 180
OY	182 CACTGCTTGGGCGCTTGTCTTCGAGAGGACATTCCGGAAGAGTCCCTCACTGTCGCA 241
DB	181 CACTGCTTGGGCGCTTGTCTTCGAGAGGACATTCCGGAAGAGTCCCTCACTGTCGCA 240
OY	242 GCTGCTGCGCCCGCCAGAGCCCAACCGGAGCCCGCCAGAGCCCGCCCTCAGAGATGA 301
DB	241 GCTGCTGCGCCCGCCAGAGCCCAACCGGAGCCCGCCAGAGCCCGCCCTCAGAGATGA 300
OY	302 TGGAGCGATGGGCTTTCTTGGCAAGAAGCGCCCAAGATGACACAGCGCGACCGGGAGG 361
DB	301 TGGAGCGATGGGCTTTCTTGGCAAGAAGCGCCCAAGATGACACAGCGCGACCGGGAGG 360
OY	362 ACAGCGAGAGAGAAAGGTTCACCTTGGCCGAGACAGGTAAACCGGGAAAGCCAGACCAAGG 421
DB	361 ACAGCGAGAGAGAAAGGTTCACCTTGGCCGAGACAGGTAAACCGGGAAAGCCAGACCAAGG 420
OY	422 GCAAAAGCCGGGGCCCATTTGGGGGGGCTGAGCCCGCTGGAGCCCAAGGGGGTCAACGGTACC 481
DB	421 GCAAAAGCCGGGGCCCATTTGGGGGGGCTGAGCCCGCTGGAGCCCAAGGGGGTCAACGGTACC 480
OY	482 CCGGGAAGCATGTGACACACAGGCAAGAGGGGCCCAAGGGCAAGAAAGGGAGCCAGAGCC 541
DB	481 CCGGGAAGCATGTGACACACAGGCAAGAGGGGCCCAAGGGCAAGAAAGGGAGCCAGAGCC 540
OY	542 TCCCAAGGCCCTCGACGCTGTGGCAGTGGCCATACCAAGTCAAGCTTTCTCGGTGCAGTGA 601
DB	541 TCCCAAGGCCCTCGACGCTGTGGCAGTGGCCATACCAAGTCAAGCTTTCTCGGTGCAGTGA 599
OY	602 CCAAGAGCTAACCCACGG 618
DB	600 CCAAGAGCTAACCCACGG 616
RESULT 7	
LOCUS	BI489428/c 708 bp mRNA linear EST 28-AUG-2001
DEFINITION	603020916T1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:519189 3',
ACCESSION	BI489428
VERSION	BI489428.1 GI:15328656
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE	NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL	1 (bases 1 to 708)
COMMENT	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgaabp-remail.nih.gov
	Tissue Procurement: Life Technologies, Inc.
	cDNA Library Preparation: Life Technologies, Inc.
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)





Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
Db	303	TATATCCGGGCAAGAGGAGACTTAAGGCGAAGAAAGGGAACTGGGCTCCAGAGCCCTG	362										
Qy	555	CAGCTGTGGCAGTGGCCATACCAGTCAAGCTTTCTCGGTGGCAGTGCACCAAGAGCTACCC	614										
Db	363	TAGCTGCGGAGTACCCAGCCAGCCAGTCCGCTTTTCGGTGGCGGTACCAAGAGTATACC	422										
Qy	615	ACGGAGGGGCTGCCATCAAGTTTGAAGAATCTGAATGAACGAGGGTGGCCACTACA	674										
Db	423	ACGTGAGGAGCTGCCCATCAAGTTTGAAGAATCTGAATGAATGAGGAGGCCACTACA	482										
Qy	675	TGCTTCCAGCGGCAAGTTCGTCTGCGGCGTGCCTGGAGTCTTACCTACCTACGACAT	734										
Db	483	TGCATCCAGTGGCAAGTTCGTCTGCGAGGCTGCACAGGAGATCTTATTACTTACCTATGACAT	542										
Qy	735	CACGCTGGCCCAACAGACCTGGCCATGGGCTGGTGGACAAACGGCCAGTACCCGATCG	794										
Db	543	TACGCTGGCCCAACAGACCTGGCCATGGGCTGGTGGACAAATGGCCAGTACCCGATTCG	602										
Qy	795	GACCTTTGATGCCAACACCGGCAACAGATGTGGCCCTCAGAGCTTCAACATCTTGCTCT	854										
Db	603	GACCTTTGATGCCAACACCGGCAACAGATGTGGCCCTCAGAGCTTCAACATCTTGCTCT	662										
Qy	855	CAAGCAGGCTGACGAAGTTTGGCTGCAATCTTCTTCACTCAAGCAGAACCG	905										
Db	663	CAAGCAGGCTGATGAGTCTGGTTACAGATTTTCTACTCGAGCAGAAATGG	713										
RESULT 9	A1916779/c												
LOCUS	A1916779	544 bp	mRNA	linear	EST 17-DEC-1999								
DEFINITION	wb26e01.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2306808 3'												
	similar to SW:ACG3 HUMAN Q15848 30 XD ADIPOCYTE COMPLEMENT-RELATED												
	PROTEIN PRECURSOR ; mRNA sequence.												
	A1916779												
ACCESSION	A1916779.1	GI:5636634											
VERSION	EST.												
KEYWORDS	Homo sapiens (human)												
SOURCE	Homo sapiens												
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;												
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.												
REFERENCE	1 (bases 1 to 544)												
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.												
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),												
JOURNAL	Tumor Gene Index												
COMMENT	Unpublished (1997)												
	Contact: Robert Strausberg, Ph.D.												
	Email: cgapbs-remail.nih.gov												
	Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael												
	R. Emmert-Buck, M.D., Ph.D.												
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima												
	Bonaldo, Ph.D.												
	cDNA Library Arrayed by: Greg Lennon, Ph.D.												
	DNA Sequencing by: Washington University Genome Sequencing Center												
	Clone distribution: NCI-CGAP clone distribution information can be												
	found through the I.M.A.G.E. Consortium												

this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonelids 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo. "

Query Match	46.1%;	Score 535.4;	DB 1;	Length 544;
Best Local Similarity	99.8%;	Pred. No. 1.9e-118;		
Matches 536;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

[illegible]



```

source
1..535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="IMAGE:281919"
/lab_host="DH10B"
/clone.lib="Soares NPL T GBC S1"
/notes="Organ: pooled; Vector: pUT73D-pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NDH1.9w, testis NH7, and B-cell
NCI GMP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "
```

Query Match 45.1%; Score 523.8; DB 2; Length 535;  
 Best Local Similarity 98.7%; Pred. No. 1.2e-115;  
 Matches 528; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	625	CTGCCCATCAAGTTTGACAAAGATTCTGTATGAACGAGGGTGGCCACTTACAAATGCTTCCAGC	684
DB	535	CTGCCCATCAAGTTTGACAAAGCTCTGTATGAACGAGGGTGGCCACTTACAAATGCTTCCATTC	476
QY	685	GGCAAGTTCGCTGGGGGCTGGCTGGGATCTACTACTTCACTTACCTAAGATCAAGCTGGACC	744
DB	475	GCCAAATTGCTGTGGGGGTGCTGGATCTGATCTACTTCACTTACCTAAGATCAAGCTGGACC	416
QY	745	AACAAGCACTGGCCATCGGCTGGTGCAACCGCCAGTACCGCATCCGACCTTTGAT	804
DB	415	AACAAGCACTGGCCATCGGCTGGTGCAACCGCCAGTACCGCATCCGACCTTTGAT	356
QY	805	GCCAAACATCGGCAACCAAGATGTGGCTTCAGGCTTCACCAATCTTGGCTCTCAAGAGGCT	864
DB	355	GCCAAACATCGGCAACCAAGATGTGGCTTCAGGCTTCACCAATCTTGGCTCTCAAGAGAGGCT	296
QY	865	GACGAAGTTGGCTGCAGATCTTCTACTCAGAGGAAACGGGCTCTTCTATGACCTTAC	924
DB	295	GACGAAGTTGGCTGCAGATCTTCTACTCAGAGGAAACGGGCTCTTCTATGACCTTAC	236
QY	925	TGCAAGACAGGCTCTTACGGGCTCTCTTAATCTATGTGCGACAGAGATGACCCCAACGAG	984
DB	235	TGCAAGACAGGCTCTTACGGGCTCTCTTAATCTATGTGCGACAGAGATGACCCCAACGAG	176
QY	985	GTATAGACATGCAAGCGGCGGTCTCTCAGAGGAAACAAAGCTTCTGACCTTGGGCTTACA	1044
DB	175	GTATAGACATGCAAGCGGCGGTCTCTCAGAGGAAACAAAGCTTCTGACCTTGGGCTTACA	116
QY	1045	GAGCAAGACCCCAACAATGTAGGCTGGGGGTGGGGGGTGCAGTAGCCGTTCTAGCTTCA	1104
DB	115	GAGCAAGACCCCAACAATGTAGGCTGGGGGTGGGGGGTGCAGTAGCCGTTCTAGCTTCA	56
QY	1105	GGCTCACTCCCTCGGCTCTTTTCTTCCCTTCATTAATCCAAACCTTTTATT	1159
DB	55	GGCTCACTCCCTCGGCTCTTTTCTTCCCTTCATTAATCCAAACATATTATT	1

RESULT 11  
 B1822472 659 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603038061.1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5179055 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1822472  
 VERSION B1822472.1 GI:15934022  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo. 1 (bases 1 to 659)	NIH-MGC	http://mgc.nci.nih.gov/.	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
Contact: Robert Strausberg, Ph.D.	Email: cga@bs-remail.nih.gov	Tissue Procurement: Life Technologies, Inc.	cDNA Library Preparation: Life Technologies, Inc.	CNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:	http://image.llnl.gov	Plate: LLNL1446	row: 1	column: 16
High quality sequence stop: 659.	Location/Qualifiers	1..659		
/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"	/clone="IMAGE:5179095"	/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."				
ORIGIN				
Query Match	43.1%	Score 500.6	DB 4	Length 659;
Best Local Similarity	95.0%	Pred. No.5.1e-110;	Indels 14;	Gaps 10
Matches 626;	Conservative 0;	Mismatches 19;		
20	CGACGCTTGGCCCGGCTGTCGCGGTGAGGAAGCCGGGACCGAGCCCGCG-AGAG	78		
1	CGACGCTTGGCCCGGCTGTCGCGGTGAGGAAGCCGGGACCGAGCCCGCGAGAG	60		
79	CTTCTTTGCTCCGAGCGCCCTTGACGTGGCGGGGAGCCGCGAGGGGTAAACA	138		
61	CTTCTTTGCTCCGAGCGCCCTTGACGTGGCGGGGAGCCGCGAGGGGTAAACA	120		
139	CCCTGGGCTCTCTGAGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	197		
121	CCCTGGGCTCTCTGAGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	180		
198	TGCTGCGAGGAGACTTCGGAAGAGGCTCCGCTCAAGTGTGTGAGCTGTGAGCTCC	255		
181	TGCTGCGAGGAGACTTCGGAAGAGGCTCCGCTCAAGTGTGTGAGCTGTGAGCTCC	240		
256	CAGGCGCCACCC-GGCCGCCCGAGAGGCCCGAGGCTTCAGAGATGA--TGGAGCA	312		
241	CAGGCGCCACCCCGGCGCCCGCGAGAGGCCCGAGGCTTCAGAGATGA--TGGAGCA	300		
313	GGCTTTCTCTGCGAAGACGCGCCAGATGACACGACGGCGACCGGGGGAGACGCGAG	372		
301	GGCTTTCTCTGCGAAGACGCGCCAGATGACACGACGGCGACCGGGGGAGACGCGAG	360		
373	GAAAGTCACTGGCGCGGACAGGTAAACGGGGGAAAGCCAGAGCCAA--GGGCA	431		
361	GAAAGTCACTGGCGCGGACAGGTAAACGGGGGAAAGCCAGAGCCAA--TGGAGCA	420		
432	GGCATTGGGCGGGCTGGCCCGCTGTGCGCCAA--GGGGGTCAACGGTACCCCGGAG	490		
421	GGCATTGGGCGGGCTGGCCCGCTGTGCGCCAA--GGGGGTCAACGGTACCCCGGAG	480		
491	ATGGACACCAAGGAGAGGGGCCCAAGGCGCAAGAAAGGGAGGCCAGGCTTCCAGG	550		

D<sub>b</sub> 481 ATGGCAGCACAGGCAAGAGGGGGCCCAAGGGCAGAGAGGCGAGCCAGGTCCTCCACGGCC 540

Q<sub>Y</sub> 551 CCTGGA-GCTGTGGCAGTGGCCATACCAAGTCAG--CTTTCTCGGTGGCAGTACCAAG 606

D<sub>b</sub> 541 CCTGATCTGTGGCAGTGGCCATACCGAAGTCAGACTATCTCGGTGGCAGTACCAAG 600

Q<sub>Y</sub> 607 AGCTACCCACGGGAGCGGCTGGCCCA-TCAAGTTTGAAGAATTTGATTAACGAGGGGTG 664

D<sub>b</sub> 601 AGCTACCCACGGGAGCGGCTGTCCATTCAAGTTTGAAGAATTTGATTAACGAGGGGTG 659

RESULT 12	AA732948/c
LOCUS	AA732948
DEFINITION	cg78d10.191 Soares fetal heart Nbh1919 Homo sapiens cDNA clone
IMAGE	894975 3' similar to SW-ACCS3 MUSE 060994 30 KD ADIPOCYTE
COMPLEMENT-RELATED PROTEIN PRECURSOR ;	mRNA sequence.
ACCESSTION	AA732948

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AA732948	AA732948.1	GI:27752888	Homo sapiens EST.	Homo sapiens (human)
			Homo sapiens	Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Okuyama, Y., Metaxoa, Chordata, Catarrhini, Hominoidea, Homo.  
Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Homo.  
1 (Pages 1 to 557)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,  
Katzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marrs, M.,  
Martin, J., Moore, B., Schellberg, K., Steptoe, M., Tan, E.,  
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
Washington Human EST Project  
Unpublished (1997)

TITLE	WashU-NCI human EST Project
JOURNAL	Unpublished (1997)
COMMENT	On Jan 14, 1998 this sequence version replaced gi:2754307

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel.: 314 286 1800  
Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Possible reversed clone; similarity on wrong strand  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 334.

FEATURES	Location/Qualifiers
source	1. .557

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1307229"
/db_xref="taxon:9606"
/clone="IMAGE:399475"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NMH19W"
/notes="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTGACCATCTGAAGGAGGAGCGCGCATCTTTTCTTTTCTTCTTCTT
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library was through one round of
normalization to a Cot = 5. Library constructed by
M.Felicia Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
DBH119W."

```

## ORIGIN

Query Match	42.5%;	Score 493;	DB 1;	Length 557;
Best Local Similarity	94.9%;	Pred. No. 3.4e-108;		
Matches 519;	Conservative	0;	Mismatches 27;	Indels 1;
				Gaps 1

OY	616	CGGAGCGGGCTGCCCATCAAGTTTGAAC -AAGATTGTATGAAACGAGGGGTGGCATTACA	674
Db	555	CGTAGGCGCTGCCCATCAAGTTTGACAAAGTTTGTATGAACGGGGGTGTCACTACAA	496
OY	675	TGCTTCCAGCGGCAAGTCGTCTGCGGGGTGCTGGATCTACTTCACTTACCTACGACAT	734
Db	495	ATCTTCCAGCGTCAAGTTTCGTCTGGGGGTGTGTTATCTACTTCACTTCACTTACGACAT	436
OY	735	CACGCTGGCCCAACAGACCTGGGCATGGGCTGGTGCACACGGCCAGTACCGATCCG	794
Db	435	CACCTGGCCMAACACCACTCGCATGGGCTGTTCACAACGGCAGTACCGCATCCG	376
OY	795	GACCTTGAATGCAACACCGGCAACCAAGATGTGGCTCAAGGCTCACCATCTGGGCTCT	854
Db	375	TACCTTGTATGCAACACCGGCAACCAAGATGTGGCTCAAGGCTCACCATCTGGGCTCT	316
OY	855	CAAGCAGGGTACGAAGTTTGGCTGCAGATCTTCTACTCAGAGCAAAACGGGCTTCTTA	914
Db	315	CAAGCAGGGTACGAAGTTTGGCTGCAGATCTTCTACTCAGAGCAAAACGGGCTTCTTA	256
OY	915	TGACCTTACTGGAACAGACAGGCTCTTTTACGGGCTTCTTAATCTATGCCCAGACAGATGA	974
Db	255	TGACCTTACTGGAACAGACAGGCTCTTTTACGGGCTTCTTAATCTATGCCCAGACAGATGA	196
OY	975	CCCCACGAGGTATAGACATGCAACGGGCGTCTCAAGGAGGGAAACAAGCTTTCGAACT	1033
Db	195	CCCCACGAGGTATAGACATGCAACGGGCGTCTCAAGGAGGGAAACAAGCTTTCGAACT	136
OY	1035	TGGGCTTACAGAGCAAGACCCCAACACTGTATGAGCTGGGGGTGTGGAGTGAACGGT	1094
Db	135	TGGGCTTACAGAGCAAGACCCCAACACTGTATGAGCTGGGGGTGTGGAGTGAACGGT	76
OY	1095	TCTAGGCTCAAGGCTCACTCTCCGCTCTTTTTCCTTTCATTAATCCAAACCTTT	1155
Db	75	TCAATACCTCAAGGCTCACTCTCTGCGCTTTTTCCTTTCATTAATCCAAACCTTT	16
OY	1155	TTATTTCA 1161	
Db	15	TTATTTCA 9	

RESULT 13	
B1488418	
LOCUS	B1488418
DEFINITION	60302091.F01 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191898 5', mRNA sequence.
	904 bp
	mRNA linear
	EST 28-AUG-2001

ACCESSION	BI488418
VERSION	BI488418.1
KEYWORDS	EST.
SOURCE	Homo sapiens
ORGANISM	Homo sapiens (human)

REFERENCE	
AUTHORS	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a>
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph. D.

Email: [cgabbs@emal.nih.gov](mailto:cgabbs@emal.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E.B. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E.B. Consortium/LNL at:  
<http://image.lnl.gov>  
Place: L1A11479 row: 0 column: 03  
High quality sequence stop: 859.

**FEATURES**  
**SOURCE**

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

/clon="IMAG:5191898"
/lab host="DH10B"
/clone.lib="NIH_MGC_114"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

```

```

ORIGIN
Query Match      40.1%; Score 466; DB 4; Length 904;
Best Local Similarity 77.3%; Pred. No. 1.2e-101;
Matches 782; Conservative 0; Mismatches 10; Indels 220; Gaps 9;

QY 1 GGAAGACTATGCTGGGGGCGAGCGCTGCGCGGCTGCTGCGGAGGAAAGCCGGGAC 60
DB 84 GGAAGACTATGCTGGGGGCGAGCGCTGCGCGGCTGCTGCGGAGGAAAGCCGGGAC 143
QY 61 GCGAGGCCCGCGAGAGCTTCTTGTCTCCGAGCGCCCTGAGAGTGCGGCGAGCCGCG 120
DB 144 GCGAGGCCCGCGAGAGCTTCTTGTCTCCGAGCGCCCTGAGAGTGCGGCGGCGGCGG 203
QY 121 AGGGTAACACATGATCCCTGGGAGTCTCTGCGCTGTGCTTCCCTGTGCTGTGAC 180
DB 204 AGGGTAACACATGATCCCTGGGAGTCTCTGCGCTGTGCTTCCCTGTGCTGTGAC 263
QY 181 CCACTGCTGGGCGCTTGTCTGCGAGGAGACTCCGGAAGGCTCCCTCACTGCTGAC 240
DB 264 CCACTGCTGGGCGCTTGTCTGCGAGGAGACTCCGGAAGGCTCCCTCACTGCTGAC 286
QY 241 AGCCTGCTGGGCGCCGAGGCGCACCGGCGCCCGAGAGCCCGAGGCGCTCAGAGATG 300
DB 287 AGCCTGCTGGGCGCCGAGGCGCACCGGCGCCCGAGAGCCCGAGGCGCTCAGAGATG 303
QY 301 ATGGGACGAAATGGGCTTCTCTGCGAAAGCGGCGCAAGTGAACAGCGCGGCGGCGG 360
DB 304 ATGGGACGAAATGGGCTTCTCTGCGAAAGCGGCGCAAGTGAACAGCGCGGCGGCGG 352
QY 361 GACAGCGGAGAGGAAGGTCCACTGCGCGGAGCAGGTAAACGGGGAAGCAGACCAAG 420
DB 353 GACAGCGGAGAGGAAGGTCCACTGCGCGGAGCAGGTAAACGGGGAAGCAGACCAAG 352
QY 421 GGCAGAACCCGGGCGCATTTGGGCGGGCTGCGCCCGTGCGCCCAAGGGGTTCAAGTACC 480
DB 353 GGCAGAACCCGGGCGCATTTGGGCGGGCTGCGCCCGTGCGCCCAAGGGGTTCAAGTACC 352
QY 481 CCCGGAAGATGCGCACACAGGCAAGAGGGGCCCAAGGGCAAGAA-AGGGAGACCAAG 539
DB 353 -ACGGGAAGATGCGCACACAGGCAAGAGGGGCCCAAGGGCAAGAAAGGGGAGACCAAG 411
QY 540 CTTTCCCGAGGCGGCTGCGAGCTGTGCGAGTGGCGCAATCAAGTCACTTTCTGTGGGCAAGT 599
DB 412 CTTTCCCGAGGCGGCTGCGAGCTGTGCGAGTGGCGCAATCAAGTCACTTTCTGTGGGCAAGT 471
QY 600 GACCAAGAGCTAACCCAGGAGGAGCGGCTGCCCATCAAGTTTGAAGAATTCATGAGAAAG 659
DB 472 GACCAAGAGCTAACCCAGGAGGAGCGGCTGCCCATCAAGTTTGAAGAATTCATGAGAAAG 531
QY 660 GGGTGGCCACTTACATGCTTTCAGAGCGCAAGTTGCTGCGGCGTGGCTTGAATCTACTA 719
DB 532 GGGTGGCCACTTACATGCTTTCAGAGCGCAAGTTGCTGCGGCGTGGCTTGAATCTACTA 591
QY 720 CTTTCACTAAGACATCAAGCTGCGCAACAGCACTGGCCATCGGGCTGTGGTGCAC-AAAG 778
DB 592 CTTTCACTAAGACATCAAGCTGCGCAACAGCACTGGCCATCGGGCTGTGGTGCAC-AAAG 651
QY 779 GCGAGTACCGGATCCGAGCTTTGA-TGCCAACAACCGGCAACAGATGTGGCTCAGGC 837
DB 652 GCGAGTACCGGATCCGAGCTTTGA-TGCCAACAACCGGCAACAGATGTGGCTCAGGC 711

```

```

QY 838 TCCACATCTCTGCTCTCAAGCAGAGGTGACAGATTTGGC-TGGAGATCTTCTACTGAG- 895
DB 712 TCCACATCTCTGCTCTCAAGCAGAGGTGACAGATTTGGC-TGGAGATCTTCTACTGAG 771
QY 896 -AGCAGAACGGGCTCTTCTAAGACC-TTACTGACAGACAGCTCTTTACGGGCTTCT 953
DB 772 AGCAGAACGGGCTCTTCTAAGACC-TTACTGACAGACAGCTCTTTACGGGCTTCT 831
QY 954 AATTATGCGGACCAAGAG-TGACCCCAAGAGGATATGACATGCGACGGCGG 1004
DB 832 AATTATGCGGACCAAGAGTTGACCCCAAGAGGATATGACATGCGGCGG 883

```

```

RESULT 14
BG381862      526 bp  mRNA  linear  EST 12-MAR-2001
LOCUS        296753 MARC 1Pig Sus scrofa cDNA 5', mRNA sequence.
DEFINITION   BG381862
ACCESSION    BG381862
VERSION      BG381862.1 GI:13306334
KEYWORDS     EST.
SOURCE       Sus scrofa (pig)
ORGANISM     Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE    1 (bases 1 to 526)
AUTHORS      Fahrentz,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Perlea,G., Sultana,R.,
Quackenbush,J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
JOURNAL      MEDLINE
PUBMED       12213789
COMMENT      Contact: Smith TPL
              USDA, ARS, US Meat Animal Research Center
              PO Box 166, Clay Center, NE 68933-0166, USA
              Tel: 402 762 4356
              Fax: 402 762 4390
              Email: smitht@mail.marc.usda.gov
              Single pass sequencing. Bases called and alt trimmed with phred
              v0.980904.e. Vector identified by cross_match with the -minscore 18
              and -mismatch 12 options.
              PCR primers
              FORWARD: AGAAGACGCTATGACCAT
              BACKWARD: GTTTCAGTCACGACG
              Plate: 1 row: 1 column: 12
              Seq primer: ATTAGGAGACATATAG.

```

```

FEATURES
SOURCE
1..526
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_id="MARC 1Pig"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

```

```

ORIGIN
Query Match      39.5%; Score 458.8; DB 4; Length 526;
Best Local Similarity 92.0%; Pred. No. 6.2e-100;
Matches 484; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 216 GAAAGGCTCCCTCACTGATGCTGACGCTGCGGCGCCCGAGGGCCCAACCGGCGCCCG 275
DB 1 GAAAGGATCCCTCACTGATGCTGACGCTGCGGCGCCCGAGGGCCCAACCTGCGCCCG 60
QY 276 AGAGCCCCAGGGGCGCTCAGAGATGATGAGACGATGGGCTTTCTGCGAAAGACGCGCA 335
DB 61 AGAGCCCCAGGGGCGCTCAGAGATGATGAGAGATGGGCTTTCTGCGAGAGATGGCGCA 120

```

QY 336 AGATGACACGACGCGACCGGCGGAGACGACGAGAGAGAGATCCACTGCGCGACAG 395  
Db 121 GGAATGGCCAGGACCGGAGACCGGCGGAGACGCGAGAGAGATCCACTGCGCGACAG 180  
QY 396 TAACCGGGGAAAGCCAGACCAAGGCGAAAGCCGGGCGCATTTGGGCGGCTGGCCCCG 455  
Db 181 TAAACCGGGGAAAGCCAGACCAAGGCGAAAGCCGGGCGCATTTGGGCGGCTGGCCCCG 240  
QY 456 TGGCGCCCAAGGGGGGTCAACGGTACCCCGGGAACATGGGACACAGGCAAGAAAGGGGCC 515  
Db 241 TGGCGCCCAAGGGGGGTCAAGTGTGCCCCAGGAGACACGCGACGCGCAAGAGAGGGGCC 300  
QY 516 CAAGGGCAAGAAAGGGAGCGAGCGCTCCAGCGCCCTGACGCTGTGAGCTGTGCGCATAC 575  
Db 301 CAAGGGCAAGAAAGGGAGCGAGCGCTCCAGAGACCTGTGACGCTGTGGCAATGCG 360  
QY 576 CAAGTCACTTTCTCGGTGCGATGACCAAGAGCTACCCAGGAGCGGCTGCCCATCAA 635  
Db 361 CAAGTCACTTTCTCGGTGCGATGACCAAGAGCTACCCAGGAGCGGCTGCCCATCAA 420  
QY 636 GTTTGAAGAATTTCTGATGAACGAGGGGTGCGCACTAAGGCTTCCAGCGCGCAATTCGT 695  
Db 421 GTTTGAAGAATTTCTGATGAAGGAGGGGTGCGCATTAATGCGTCAATGGCAAGTTTGT 480  
QY 696 CTGCGCGGTGCTGGAGTTACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 741  
Db 481 CTGCGCGGTGCTGGAGTTACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 526

RESULT 15  
AI288625 455 bp mRNA linear EST 29-JAN-1999  
LOCUS G188625.x1 Soares\_NbHMPu\_S1 Homo sapiens cDNA clone IMAGE:1879377  
DEFINITION 3 similar to SW:ACR3 HUMAN Q15848 30 KD ADIPOCTYB  
COMPLEMENT-RELATED PROTEIN PRECURSOR ; mRNA sequence.  
ACCESSION AI288625  
VERSION AI288625.1 GI:3932128  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 455)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 1078 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 416.  
Location/Qualifiers  
1..455  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1879377"  
/tissue\_type="Pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/clone\_id="Soares\_NbHMPu\_S1"  
/note="Organ: mixed (see below); Vector: pTT73D-Pac  
(pharmacia) with a modified polylinker; Site 1: Not 1;  
Site 2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NbHMPu, pregnant uterus  
NbHMPu, and fetal heart NbHMPu) were mixed, and 88 circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries. The pools

FEATURES  
SOURCEJOURNAL  
COMMENTREFERENCE  
AUTHORS  
TITLEACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISMRESULT 15  
LOCUS  
DEFINITIONACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISMFEATURES  
SOURCE

ORIGIN  
Query Match 39.0%; Score 452.4; DB 1; Length 455;  
Best Local Similarity 99.6%; Pred. No. 2.1e-96;  
Matches 453; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."

QY 706 CTTGGATCTACTACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 765  
Db 455 CTTGGATCTACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 396  
QY 766 CTGATGACAAAGGCGGAGTACCGCATCCGACCTTTGATGACCAACCGGCAACACGAT 825  
Db 395 CTGATGACAAAGGCGGAGTACCGCATCCGACCTTTGATGACCAACCGGCAACACGAT 336  
QY 826 GTGGCTCAGGCTTCACCATCTGCTCTCAAGCAGGATGACGAAGTTTGGCTGCAATC 885  
Db 335 GTGGCTCAGGCTTCACCATCTGCTCTCAAGCAGGATGACGAAGTTTGGCTGCAATC 276  
QY 886 TTCTACTCAGAGAGAAAGGGGCTCTTCTATGACCTTACTGACAGACAGCTCTTTACG 945  
Db 275 TTCTACTCAGAGAGAAAGGGGCTCTTCTATGACCTTACTGACAGACAGCTCTTTACG 216  
QY 946 GGCCTTCTAATCTATGCGGACCAAGATGACCCCAAGGATATGACATGCCAGCGGAT 1005  
Db 215 GGCCTTCTAATCTATGCGGACCAAGATGACCCCAAGGATATGACATGCCAGCGGAT 156  
QY 1006 CTTCCAGGCAAGGAAAGAGCTTTCTGACTTGGCTTTCAGAGCAAGCCCCCAACTGTA 1065  
Db 155 CTTCCAGGCAAGGAAAGAGCTTTCTGACTTGGCTTTCAGAGCAAGCCCCCAACTGTA 96  
QY 1066 GGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1125  
Db 95 GGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 36  
QY 1126 TTTTTCCTTCAATTAATCAAACTTTTATTC 1160  
Db 35 TTTTTCCTTCAATTAATCAAACTTTTATTC 1

Search completed: December 25, 2004, 06:40:08  
Job time: 4248 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2004, 03:17:57 ; Search time 123 Seconds

(without alignments)  
6709.154 Million cell updates/sec

Title: US-10-621-787-1

Perfect score: 1161  
Sequence: 1 ggaataactagcctgagggcc.....atccaactttttatcca 1161

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgm2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgm2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgm2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgm2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgm2\_6/ptodata/1/ina/PTCUS\_COMB.seq:\*  
6: /cgm2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1161	100.0	1161	US-09-552-204A-1	Sequence 1, Appl1
2	702.2	60.5	1107	US-09-188-930-217	Sequence 217, App
3	702.2	60.5	1107	US-09-312-283C-217	Sequence 217, App
4	701.6	60.4	1052	US-09-312-283C-358	Sequence 358, App
5	613.6	52.9	855	US-09-552-204A-10	Sequence 10, Appl
6	453	39.0	750	US-09-188-930-19	Sequence 19, Appl
7	289.6	24.9	536	US-09-312-283C-19	Sequence 19, Appl
8	253.2	21.8	1333	US-09-552-204A-11	Sequence 11, Appl
9	101.2	8.7	729	US-09-227-357-51	Sequence 51, Appl
10	101.2	8.7	729	US-09-140-804-10	Sequence 10, Appl
11	101.2	8.7	729	US-09-686-838B-10	Sequence 10, Appl
12	96.8	8.3	4517	US-09-140-804-9	Sequence 9, Appl1
13	96.8	8.3	4517	US-09-686-838B-9	Sequence 9, Appl1
14	96.8	8.3	4517	US-09-776-976-5	Sequence 5, Appl1
15	96.8	8.3	4517	US-09-909-547-5	Sequence 5, Appl1
16	96.8	8.3	4545	US-09-569-852B-5	Sequence 5, Appl1
17	95.8	8.3	1313	US-08-463-911-6	Sequence 6, Appl1
18	91.6	7.9	1347	US-09-140-804-1	Sequence 1, Appl1
19	91.6	7.9	1347	US-09-686-838B-1	Sequence 1, Appl1
20	91.6	7.9	1377	US-09-866-028-41	Sequence 41, Appl1
21	91.6	7.9	1377	US-10-140-002-361	Sequence 361, App
22	91.6	7.9	1377	US-09-944-457-41	Sequence 41, Appl
23	89.2	7.7	20966	US-09-776-976-7	Sequence 7, Appl1
24	89.2	7.7	20966	US-09-909-547-7	Sequence 7, Appl1
25	89.2	7.7	20966	US-09-569-852B-1	Sequence 1, Appl1
26	87.2	7.5	1276	US-08-463-911-1	Sequence 1, Appl1
27	87.2	7.5	1276	US-09-776-976-3	Sequence 3, Appl1

28	87.2	7.5	1276	4	US-09-909-547-3	Sequence 3, Appl1
29	84.6	7.3	728	4	US-09-336-536-2	Sequence 2, Appl1
30	84.6	7.3	1338	4	US-09-336-536-1	Sequence 1, Appl1
31	83.4	7.2	2543	1	US-08-555-669-11	Sequence 11, Appl
32	83.4	7.2	2543	3	US-09-073-663-11	Sequence 11, Appl
33	81.4	7.0	1152	4	US-09-776-976-1	Sequence 1, Appl1
34	81.4	7.0	1152	4	US-09-909-547-1	Sequence 1, Appl1
35	80.6	6.9	4167	4	US-09-169-768-12	Sequence 12, Appl
36	79	6.8	660	4	US-09-169-768-43	Sequence 43, Appl
37	79	6.8	3170	4	US-09-169-768-1	Sequence 1, Appl1
38	79	6.8	3171	4	US-09-169-768-15	Sequence 15, Appl
39	79	6.8	3171	4	US-09-169-768-19	Sequence 19, Appl
40	79	6.8	3181	1	US-08-655-086-1	Sequence 1, Appl1
41	79	6.8	3349	4	US-09-169-768-13	Sequence 13, Appl
42	79	6.8	3531	4	US-09-169-768-7	Sequence 7, Appl1
43	79	6.8	3541	4	US-09-169-768-9	Sequence 9, Appl1
44	79	6.8	4409	4	US-09-331-347C-22	Sequence 22, Appl
45	78.4	6.8	1839	1	US-08-383-744-1	Sequence 1, Appl1

#### ALIGNMENTS

```
RESULT 1
US-09-552-204A-1
; Sequence 1, Application US/09552204A
; Patent No. 6620909
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACRP2
; FILE REFERENCE: 99-08
; CURRENT APPLICATION NUMBER: US/09/552,204A
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,207
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(987)
; US-09-552-204A-1

Query Match      100.0%; Score 1161; DB 4; Length 1161;
Best Local Similarity 100.0%; Pred. No. 8.4e-289;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAATACTATGCTGGGCGGCGAGCGCTGTGCGGCTGTGCGGCTGTGAGGAAAGCCGGGAG 60
DB      1 GGAATACTATGCTGGGCGGCGAGCGCTGTGCGGCTGTGCGGCTGTGAGGAAAGCCGGGAG 60
QY      61 GCGAGAGCCCGCCGAGAGGCTTTTGTCTCGGAGAGCCCTTGAGAGCTGGCGGCGGCGG 120
DB      61 GCGAGAGCCCGCCGAGAGGCTTTTGTCTCGGAGAGCCCTTGAGAGCTGGCGGCGGCGG 120
QY      121 AGGTTAACACCATGATCCCTGGGCTGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 180
DB      121 AGGTTAACACCATGATCCCTGGGCTGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 180
QY      181 CCACTGCTTGCGGCGCTTGTGCTGCGAGGAGCTTTCGGAAGGCTCCCTCAACTGCTGTGC 240
DB      181 CCACTGCTTGCGGCGCTTGTGCTGCGAGGAGCTTTCGGAAGGCTCCCTCAACTGCTGTGC 240
QY      241 AGCTTGCTTGCGGCGCTTGTGCTGCGAGGAGCTTTCGGAAGGCTCCCTCAACTGCTGTGC 300
DB      241 AGCTTGCTTGCGGCGCTTGTGCTGCGAGGAGCTTTCGGAAGGCTCCCTCAACTGCTGTGC 300
QY      301 ATGGAGCAATGGGCTTCTCTGCAAGAGCGGCAATGAGCAAGAGCGGCGGCGG 360
DB      301 ATGGAGCAATGGGCTTCTCTGCAAGAGCGGCAATGAGCAAGAGCGGCGGCGGCGG 360
```

Db	301	ATGGAGCAATATGGGCTTTCTGTGGCAAGAACGGCCAAAGATGACAACAACGGGAGACCGGGGG	360
QY	361	GACACCGGAGAGGAAAGTTCACCTCTGGCTCGGACAAGTTAAACCGGGGAAAGCCACAGACCAAAAG	420
Db	361	GACACCGGAGAGGAAAGTTCACCTCTGGCTCGGACAAGTTAAACCGGGGAAAGCCACAGACCAAAAG	420
QY	421	GGCAAAAGCCGGGGGCATTGGGGCGGGCTGGCCCGCTGGACCCCAAGGGGGGTCAACGGTATCC	480
Db	421	GGCAAAAGCCGGGGGCATTGGGGCGGGCTGGCCCGCTGGACCCCAAGGGGGGTCAACGGTATCC	480
QY	481	CCCGGGAAGCATGGGCACACAGGACAAAGAGGGGCCAAAGGGCCAAAGAAAGGGAGGCCAGGC	540
Db	481	CCCGGGAAGCATGGGCACACAGGACAAAGAGGGGCCAAAGGGCCAAAGAAAGGGAGGCCAGGC	540
QY	541	CTCCACGGGCCCCCTGCAGCTGTGTGGAGAGTGGCCATACCAAGTACGCTTCTCGGTGGCAATGG	600
Db	541	CTCCACGGGCCCCCTGCAGCTGTGTGGAGAGTGGCCATACCAAGTACGCTTCTCGGTGGCAATGG	600
QY	601	ACCAAGAGCTACCCACCGGAGCGGGCTGGCCATCAAGTTTGAACAAGATTCTGATGAACGAG	660
Db	601	ACCAAGAGCTACCCACCGGAGCGGGCTGGCCATCAAGTTTGAACAAGATTCTGATGAACGAG	660
QY	661	GGTGGCCACTACATGCTTCTTCCAGCGGGCAAGTTTGTCTTGGGGCTGTGGCTTGGGATCTTACAC	720
Db	661	GGTGGCCACTACATGCTTCTTCCAGCGGGCAAGTTTGTCTTGGGGCTGTGGCTTGGGATCTTACAC	720
QY	721	TTTCACTTACGACATCAAGCTGAGCCCAAACAAGACCTGGGCTATCGGCTTGGTGAACAACGGC	780
Db	721	TTTCACTTACGACATCAAGCTGAGCCCAAACAAGACCTGGGCTATCGGCTTGGTGAACAACGGC	780
QY	781	CAGTACCGGATCCGGACCTTTGATGATGCCAAACCCGGACAACAAGATGTGGCTTCAGGCTCC	840
Db	781	CAGTACCGGATCCGGACCTTTGATGATGCCAAACCCGGACAACAAGATGTGGCTTCAGGCTCC	840
QY	841	ACCAATCCGGGCTTCACAGCAGGGTGAACAATTTGGCTGCGAGATTTTACCTCAGAGGAG	900
Db	841	ACCAATCCGGGCTTCACAGCAGGGTGAACAATTTGGCTGCGAGATTTTACCTCAGAGGAG	900
QY	901	AACGGGCTCTTCTATGACCTTTACTGACACAGACCTTTTACGGGCTTCTTAATCTAT	960
Db	901	AACGGGCTCTTCTATGACCTTTACTGACACAGACCTTTTACGGGCTTCTTAATCTAT	960
QY	961	GCCGACCAAGATGACCCCAACGAGGTATGACATGCCACGGGGGTCTTCCACGACAGGGAA	1020
Db	961	GCCGACCAAGATGACCCCAACGAGGTATGACATGCCACGGGGGTCTTCCACGACAGGGAA	1020
QY	1021	CAAGCTTCTGACATTGGGGCTTACAGAGCAAGACCCCAACATGTATGGCTGGGGGGTGGGGG	1080
Db	1021	CAAGCTTCTGACATTGGGGCTTACAGAGCAAGACCCCAACATGTATGGCTGGGGGGTGGGGG	1080
QY	1081	GTCGAGTAGAGGGTTCTAGCCTCAAGGCTCAACCTCTCGGCTCTTTTTCCTCCCTTCATT	1140
Db	1081	GTCGAGTAGAGGGTTCTAGCCTCAAGGCTCAACCTCTCGGCTCTTTTTCCTCCCTTCATT	1140
QY	1141	AAATCCAAACCTTTTATTTCA	1161
Db	1141	AAATCCAAACCTTTTATTTCA	1161

RESULT 2  
 US-09-188-930-217  
 Sequence 217, Application US/09188930A  
 Patent No. 6150502  
 GENERAL INFORMATION:  
 APPLICANT: Watson, James D.  
 APPLICANT: Strachan, Iorna  
 APPLICANT: Sleeman, Matthew  
 APPLICANT: Ornstut, Rene  
 APPLICANT: Murison, James Greg  
 TITLE OR INVENTION: Compositions Isolated From Skin Cells  
 TITLE OR INVENTION: and Methods For Their Use  
 FILE REFERENCE: 11000.101cl  
 CURRENT APPLICATION NUMBER: US/09/188,930A

```

: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 217
: LENGTH: 1107
: TYPE: DNA
: ORGANISM: Rat
US-09-188-930-217

```

Query Match	60.5%;	Score 702.2;	DB 3;	Length 1107;
Best Local Similarity	84.7%;	Pred. No. 6.1e-171;		
Matches 800; Conservative	0;	Mismatches 143;	Indels 2;	Gaps 1;

QY	121	AGGGTAACCAACCAAGATACCCCTGGATGTCTCTGGACCTGGCCCTCCCTGGCTGTCTAC	180
Db	162	AAGGTGACCAACCAAGATCTCCTGGATGTCTTGGCTGTGGCCCTTCGGTGTCTGTAC	221
QY	181	CCACTGCTTGGCGCCTTTGCTCGACGGGACCTTCGGAAAAGCTCCCTCTCAACTGTCTGC	240
Db	222	CCAAATGCTTGTGTGCTTGTGTCTCGAAGGGACCTTCAGAAAGGTGTGTCTCAACTGTGTGGC	281
QY	241	AGCTGCGCTGGCCCCCAAGGGCCCAACCGGAGCCCCCAAGAGGCCCCAGAGGCGCTCAGAAATG	300
Db	282	AGTCTGCTGTGTCCCAAGGCGCCACTGTGGCCCTTCAGAGACCAACAGATCTCTCAGAAATG	341
QY	301	ATYGGACGAATYGGGCTTTCTGTGCAAAAGACGGCCAAAGATGGAACAACAAGCGGCGACCGGGGG	360
Db	342	GTGGGAAAGAAATGGGGTTTTCTGTGTAAGGATGGCCAAAGCGGCGCAGACCGAGACCGAAGG	401
QY	361	GACAGCGGAGAGAAAGGTCCAACCTGGCCGGAACAAGTAAACCGGGGAAABCCAGAGACCAAG	420
Db	402	GACAGTGAAGAAAGGTCCAACCTGGCAAGACAGGCAACCAAGGAAAACAAGAACCAAG	461
QY	421	GGCAAAAGCCGGGGGCAATTTGGGGCGGGCTGGCCCCCGTGGCCCCAAAGGGGGTCAACGGTAAC	480
Db	462	GGCAAAAGCTGGGGGCAATTTGGGAGAGCGGGTCTTGAAGGACCAAGAGGGGGTCAAGTGTATCC	521
QY	481	CCCGGAAAGCATGGGCAACAACAGGCAAGAGGGGCCCAAGGCGCAAGAAAGGGGAGCCAGGC	540
Db	522	CCCCGAAACATGTGTATACCCGGGCAAGAGGGGACCTTAAGGGCAAGAAAGGGGAACTTGGG	581
QY	541	CTCCAGGGCCCCCTGACGTGTGGAGAGTGGCCATACCAAGTCAGACTTTCTCGGTGGCAGTG	600
Db	582	CTCCAGGGCCCCCTGTAGCTGTGGGAGATACCGAGACCAAGTGGGCTTTTCGGTGGCGGTA	641
QY	601	ACCAAGAGCTACCCAAGGGAGCGGCTGGCCATCAAGTTTGAACAAGTTCTGATGAACGAG	660
Db	642	ACCAAGAGTTATCCACGAGGACGACTGGCCATCAAGTTTGAACAAGTTCTGATGAATAG	701
QY	661	GGTGGCCCATTAATGTCTTCAAGGGGGAAGTTTCGTGGGGGCGTGGCCCTGGGATCTAATAC	720
Db	702	GGAGGCCCATTAACAATGCAATCGAGTGGCAAGTTTCGTGCACGCTGSCCAAGGAAATCTAATAC	761
QY	721	TTCACTTAACAATACGCTGGGCCCAACAAGCACTGGCCATTCGGCTGTGTGCAACAACGGC	780
Db	762	TTTATCTAATGACATTAAGCTGGCCCAACAACAACCTGGCCATTCGGCTGTGTGCAACAATGGC	821
QY	781	CAGTACCGCATCCGGACCTTTGATGGCCAAACCGGCAACAACAGATGTGGGCTCAGGCTCC	840
Db	822	CAGTACCGCATTCGGACTTTTGAAGCCCAACAACCGGCAACAACAGATGTGGGCTCAGGCTCC	881
QY	841	ACCAATCCGTGCTCAAGCAGAGGTGACGAAGTTTGGCTGCAAGATCTTACTCAGAGCAG	900
Db	882	ACCAATCCGTGCTCAAGAGAGGGGTGATGAATCTGGTTACAGATTTTCTACTCGAGACAG	941
QY	901	AACGGGCTTTCTATGAACCTTACTGGAACAAGCCTTTTACGGGCTTCTAATCTAT	960
Db	942	AATGGACTTTCTATGAACCTTTATTGGAACGACAGCCTGTTCACCGGCTTCTCTCATCTAC	1001
QY	961	GCCGACCAAGATACCCCAACGAAGTATAGCAATGCCAGG--CGGTCTTCAAGGCAAGG	1011
Db	1002	GCTATATCAAGGAGACCCCAATGAAGGTATAGCAAGCTGGGGTTGAGCGGTTCAGGACGAGG	1061



```
QY 421 GGCAGAACCCGCGGCATTTGGCGGCTGGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACC 480
| | | | |
Db 333 GGCAGAACCTGGGCGCATTTGGGAGAGGGGGTCTCGAGGAGCCCAAGGGGGTCACTGGTACC 392
| | | | |
QY 481 CCCGGGAGCATGGCACAACGAGGAGAGGGGGCCCAAGGGGCAAGAAAGGGGAGCCAGGC 540
| | | | |
Db 393 CCCGGGAGCATGGTGTATACGGGGCAAGAGGAGGCTTAAGGGCAAGAAAGGGGAACTGGGG 452
| | | | |
QY 541 CTCGCAAGGCCCCGCGAGCTGTGGCAAGTGGCCATACCAAGTCACTTCTCGTGGGAGTG 600
| | | | |
Db 453 CTCGCAAGGCCCCCTGTAGCTGTGGCGAGTGGCCAGCCAGTGGCCCTTTCTGGTGGGTA 512
| | | | |
QY 601 ACCAAGAGCTACCCAGCGGAGCGGCTGCCATCAAGTTTGAACAAGATTCTGATGAAGAG 660
| | | | |
Db 513 ACCAAGAGTTACCAAGTGAAGCACTGCCATCAAGTTTGAACAAGATTCTGATGAAGAG 572
| | | | |
QY 661 GGTGGCACTACATGCTTTCAGCGGCAAGTGTCTGCGGGCTGCTGGGATCTACTAC 720
| | | | |
Db 573 GGAAGCCACTACATGATCATCAGTGGCGAGTGTCTGCGAGCGTGGCCAGGATCTATTAC 632
| | | | |
QY 721 TTCACTACGACATCAAGCTGGCCAAAGAGCACTGGCCATCGGCTGTGGTCAACAGCGC 780
| | | | |
Db 633 TTTCACCTATGACATTAAGCTGGCCAAACACCTGGCCATCGGCTGTGGTCAACATAGGC 692
| | | | |
QY 781 CAGTACCGCATCCGACCTTTGATGCCCAACACCGGCAACGATGTGGCTCAGGCTCC 840
| | | | |
Db 693 CAGTACCGCATTCGCACTTTTGAAGCCCAACACCGGCAACGATGTGGCTCAGGCTCC 752
| | | | |
QY 841 ACCATCTGCTCTCAAGCAGGGGTGACGAAGTTTGGCTGCAATCTTCTACTCAAGCAG 900
| | | | |
Db 753 ACCATCTGCTCTCAAGGAGGGGTGATGAGTCTGGTTTCAAGATTTTCTACTCGAGCAG 812
| | | | |
QY 901 AACGGGCTCTTCTATGACCTTTACTGAGACAGACCTCTTTTACGGGCTTCTAATCTAT 960
| | | | |
Db 813 AATGACTCTTCTTCAACACCTTTATGAGACCGACGCTGTTCACGGCTTCTCTCACTAC 872
| | | | |
QY 961 GCCGACAGAGTACCACCAAGAGTATGACATGCGCACG -CGTCTCTCAGGCAAGGA 1019
| | | | |
Db 873 GCTGATCAAGAGAACCCCAATGAGTATGACAAGCTGGGGTTGAGCGTCAAGGAGGA 932
| | | | |
QY 1020 ACAAGCTTCTGAGCTTGGGCTTTCAGAGCAAGACCCCACTGTAGGCTGGGGTGGGG 1079
| | | | |
Db 933 CTAAAGATTCGCAAGGGTGTGATGAGAGGATCTCTGAATCGAGGCTGGGCACTGGCA 992
| | | | |
QY 1080 GGTGAGTGAAGGGTCTTAGCCTCAGGCTCAGCTCCGCTCTTTT 1129
| | | | |
Db 993 GTTCTTGGAGC--TTTATTCCCAAGGCAAGCTCTCTGTGCTGCTTT 1040
| | | | |

RESULT 5
US-09-552-204A-10
; Sequence 10, Application US/09552204A
; Patent No. 6620909
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; TITLE OF INVENTION: ADIPOCTE-SPECIFIC PROTEIN HOMOLOG ZACRP2
; FILE REFERENCE: 99-08
; CURRENT APPLICATION NUMBER: US/09/552,204A
; PRIORITY FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,207
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate nucleotide sequence encoding the
; NAME/KEY: variation
; LOCATION: (1)...(855)
```

```
OTHER INFORMATION: Each N is independently any nucleotide.
US-09-552-204A-10

Query Match 52.9%; Score 613.6; DB 4; Length 855;
Best Local Similarity 60.4%; Pred. No. 3,2e-148;
Matches 516; Conservative 170; Mismatches 168; Indels 0; Gaps 0;

QY 133 ATGATCCCTGGGTGCTCTGCGCTGTGGCTCCCTCCCTGCTGCTGACCACTGCTTGGC 192
| | | | |
Db 1 ATGATTCCTTGGGTNTTNTTNGCNTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
| | | | |
QY 193 GCCTTGTCTGCAAGGACTTTCGGAAGGCTCCCTCAACTGTGTCTGCAAGCTGCTGGC 252
| | | | |
Db 61 GCTTTCGCTGCAAGGACTTTCGGAAGGCTCCCTCAACTGTGTCTGCAAGCTGCTGGC 120
| | | | |
QY 253 CCCCAGGCGCCACCCGCGCCCCCAAGAGGCCCCCAAGGCGCTCCAGGATGATGGAGCAATG 312
| | | | |
Db 121 CCNCAAGGCGCCACCCGCGCCCCCAAGAGGCCCCCAAGGCGCTCCAGGATGATGGAGCAATG 180
| | | | |
QY 313 GGCCTTCTGCGCAAGAGCGCCCAAGATGACAGAGCGCGGAGCCGCGGAGAG 372
| | | | |
Db 181 GCTTTCGCTGCAAGGACTTTCGGAAGGCTCCCTCAACTGTGTCTGCAAGCTGCTGGC 240
| | | | |
QY 373 GAAAGTCACTGCGCGGCAAGGTAACCGGGAAGCCAGAGCAAGAGGCAAGGCGG 432
| | | | |
Db 241 GARGGCGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
| | | | |
QY 433 GCCATTTGGGGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 492
| | | | |
Db 301 GCNATHTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
| | | | |
QY 493 GGCACACAGAGGCAAGAGGAGGCGCCCAAGGCGCAAGAAAGGAGGCGGCGGCGGCGGCGGCGGCGG 552
| | | | |
Db 361 GGNACACAGAGGCAAGAGGAGGCGCCCAAGGCGCAAGAAAGGAGGCGGCGGCGGCGGCGGCGGCGG 420
| | | | |
QY 553 TGCAGCTGTGGCAAGTGGCAATACCAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
| | | | |
Db 421 TGYWNTGTGCGWNSGNCAYACNARWNSGNTTWSNGTNGCNGTNAACNARWNSNTAY 480
| | | | |
QY 613 CCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 672
| | | | |
Db 481 CCNAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
| | | | |
QY 673 AATGCTTCCAGGCGGCAAGTGTCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
| | | | |
Db 541 AATGCTTCCAGGCGGCAAGTGTCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
| | | | |
QY 733 ATCAGCTGTGGCAAGGACTGCTGCGCATGCGCTGTGTGCAACAGCGGCAAGTACCGCATC 792
| | | | |
Db 601 ATCAGCTGTGGCAAGGACTGCTGCGCATGCGCTGTGTGCAACAGCGGCAAGTACCGCATC 660
| | | | |
QY 793 CGGACCTTTGAGCCCAACCCGCAACCAAGATGAGGCGCTCAGGCTCCAGCATCTGGGCT 852
| | | | |
Db 661 GGNACCTTTGAGCCCAACCCGCAACCAAGATGAGGCGCTCAGGCTCCAGCATCTGGGCT 720
| | | | |
QY 853 CTCAGCAGGCGGCAAGTGTCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
| | | | |
Db 721 YTTAARCAAGGCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
| | | | |
QY 913 TATGACCTTACTGCAAGACAGCTCTTTAGGGGCTTCTTAATCTATGCGGCAAGT 972
| | | | |
Db 781 TATGACCTTACTGCAAGACAGCTCTTTAGGGGCTTCTTAATCTATGCGGCAAGT 840
| | | | |
QY 973 GACCCCAAGAGGT 986
| | | | |
Db 841 GATCCNAAGAGGT 854
| | | | |

RESULT 6
US-09-188-930-19
; Sequence 19, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
```



```

? APPLICANT: Watson, James D.
? APPLICANT: Strachan, Lorna
? APPLICANT: Sleeman, Matthew
? APPLICANT: Onrust, Rene
? APPLICANT: Mullison, James Greg
? TITLE OF INVENTION: Compositions Isolated From Skin Cells
? TITLE OF INVENTION: and Methods For Their Use
? FILE REFERENCE: 11000.1011c1
? CURRENT APPLICATION NUMBER: US/09/188,930A
? CURRENT FILING DATE: 1998-11-09
? NUMBER OF SEQ. ID NOS: 348
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 19
? LENGTH: 750
? TYPE: DNA
? ORGANISM: Rat
US-09-188-930-19

```

Query Match	39.0%	Score 453, DB 3, Length 750,
Best Local Similarity	86.2%	Pred. No. 5, 5e-107,
Matches 501, Conservative	0,	Mismatches 80, Indels 0, Gaps 0

QY	121	GGGGTAAACCAACATATATCCCTGGGGGCTCTGGACCTGGACCCTCCCTGGTGTCTGAC	180
Db	170	AAGGTGACCAACATATATCTCTGGATGCTCTTGGCTGTGACCTTCCGTGTGCTGTAC	229
QY	181	CCACTGCTTGGCGCCTTGTCTGTGCAGGGACTTCCGAAAGGCTCCCTCAACTGTGTGC	240
Db	230	CCAAATGCTTGGGCTTGTGTCTGTGCAGGGACTTCCAGAAAGGTGTGTCTCAACTGTGTGC	289
QY	241	AGCTGCTTGGCCCCCAGGAGCCCAACCCGGCCCCCAGGAGCCCTAAGGACTTAAGAAATG	300
Db	290	AGTCTGCTGTGTCCCAAGGCCCACTGGCCCTCCAGGAGCACCAAGATCTTCAGAAATG	349
QY	301	ATYGGAGCAATATGGGCTTTCTCTGGCAAAAGACGGCCAAATATGAGACACGACGGACCGAGGG	360
Db	350	GTGGGAAGAAATGGGTTTCTCTGTAAAGATGACCAGACAGGCCAGGACCGAAGACCCAGGG	409
QY	361	GACACGGAGAGGAAGGTCCACTTGGCCGGACAGGTAACTCGGGGAAAGCCAGGACCAAAAG	420
Db	410	GACAGTGAAGAAAGAGTCCACTGTGGCAGGACAGGCAACCGAAGAAACAAAGACCAAAAG	469
QY	421	GGCAAAAGCCGGGGCCATTGGGCGGGGCTGGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACC	480
Db	470	GGCAAAAGCTGGGGCCATTGGGAGACCGGGTCTTCAGAGAACCAAGGGGGGTCAATGTGTACC	529
QY	481	CCCGGGAAGCATATGGCACACAGGCAAGAGGGGGCCCAAGGGCAAGAAAGGGAGCCAGGC	540
Db	530	CCCGGGAATCATGTGTATATCCGGGCAAGAGGGGACTTAAGGGCAAGAAAGGGAAACCTGGG	589
QY	541	CTCCAGGCCCCCTGACGCTGTGGCAGTGGCACTTACCAATCAGCTTTCTCGTGGCAGTG	600
Db	590	CTCCAGGCCCCCTGTAGCTGTGGGCAGTATCCAGGCCAAATCGGCTTTTTCGGTGGCGGTA	649
QY	601	ACCAAGAGCTATCCACGGGAGCGGCTGCCCATCAAGTTTGAACAAGTTCTGATGAACGAG	660
Db	650	ACCAAGAGTTTCCCACTGAGGCACTGCCCATCAAGTTTGAACAAGATTCTGATGAATGAG	709
QY	661	GGTGGCACTCAATGCTTCCACGGGCAAGTTCGTGTGGG	701
Db	710	GGAGGCCACTACATATGCATCCAGTGGCAAGTTGTGTGAG	750

RESULT 7  
US-09-312-283C-19  
Sequence 19, Application US/09312283C  
Patent No. 6573095  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onnisc, Rene  
APPLICANT: Mutison, James G.

```

1  APPLICANT: Kimble, Krishanand D.
2  TITLE OF INVENTION: Compositions Isolated from Skin Cells
3  TITLE OF INVENTION: and Methods for Their Use
4  PILE REFERENCE: 11000.1011c2
5  CURRENT APPLICATION NUMBER: US/09/312,283C
6  CURRENT FILING DATE: 1999-05-14
7  NUMBER OF SEQ ID NOS: 425
8  SOFTWARE: FastSeq for Windows Version 4.0
9  SEQ ID NO 19
10 LENGTH: 750
11 TYPE: DNA
12 ORGANISM: Mouse
13 US-09-312-283C-19
14
15 Query Match      39.0%; Score 453; DB 4; Length
16 Best Local Similarity 86.2%; Pred. No. 5,5e-107;
17 Matches 501; Conservative 0; Mismatches 80; Indel

```

QY 121 AGGGTAAACCAACATGATCCCTGGAGTCTGGCCTGGCCCTCCCTGCTGCTGAC 180  
Db 170 AAGGTGACCAACATGATCTCTGGATGCTTGGCCTGGGCCCTTCCGATGCTGCTGAC 229

Oy 181 CCACGTGTGACGCGCTTTGCTGCGAGGACCTTCCGGAAGAGCTCCCTCACTGAGTCTGC 240  
 Db 230 CCAATGCTTGGTGCTTTTGCTGCGAGGACCTTCGAAGAGGTGATCTCACTGGTGTGC 289  
 Oy 241 AGCTGCTTGCGCCCTCCAGGAGCCCAATCCGGGCCCCCGAGAGCCCGAGGCGCTTCAGGAATG 300  
 Db 290 AGTCTGCTGTGATCCCAAGGCGCAACTGGCCCTCCAGGAGGACACAGGATCTTCAGGAATG 349  
 Oy 301 ATGGGACCAATGAGGCTTTCTCTGGCAAAAGACGCGCCAAAGATGACACAGACGGGACCCGAGG 360  
 Db 350 GTGGGAAGAAATGGGTTTTCTGTGTAAAGATGGCCAAAGACGGCCAGGACGGAGACCGAGGG 409  
 Oy 361 GACAGCGGAGAGGAAGGTTCACCTTGGCCGGACAGGTAAATCCGGGGAAGACCGAGACCAAG 420  
 Db 410 GACAGTGAAGAGAAAGAGTTCACCTTGGCAGGACAGGCAACCGAGGAAAAACAGGACCAAG 469  
 Oy 421 GGGAAAGCCGGGAGCCATTGGGCGGCGCTGGCCCCCGCTGGCGCCCAAGGGGAGTCAACGTTACC 480  
 Db 470 GGGAAAGCTGGGGGCCATTGGGAGAGACCGGGTCTTCAGAGACCCAAAGGGGAGTCAAGTGTACC 529  
 Oy 481 CCGGGAAGCATGTCACACACAGGCAAGAAAGGGGCCCAAGGGCCAGAAAGGGAGCCAGGC 540  
 Db 530 CCGGGAAGCATGGTATACCGGGCAAGAAAGGACCTTAAGGGGCAAGAAAGGGGAAACTGGG 589  
 Oy 541 CTCCAGGCGCCCTGAGCGTGTGGCGAGTGGCCATTAACATGACGTTTCTCGGTGGCAGTG 600  
 Db 550 CTCCAGGCGCCCTGTAGCTGTGGCGAGTGAAGCCAGGCAAGTCCGGCTTTTTCGGTGGCGGTA 649  
 Oy 601 ACCAAGAGCTTACCCACGGGAGCGGCTGCCCATCAAGTTTGACAAAGATTCTGATGAACGAG 666  
 Db 650 ACCAAGAGATTACCCACGTGAGCGGACGTGCCCATCAAGTTTGAACAAAGATTCTGATGAATGAG 709  
 Oy 661 GGTGGCACTACAATGCTTCCACGAGCGCAAGATTGATGCGG 701  
 Db 710 GGAAGCCACTCAATGCAATCCAGTGGCAAGTTTGCTGTGAG 750

```

RESULT 8
US-09-552-204A-11
; Sequence 11, Application US/09552204A
; Patent No. 6620909
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACRP2
; FILE REFERENCE: 99-08
; CURRENT APPLICATION NUMBER: US/09/552,204A
; CURRENT FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,207
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 20

```

```
SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(360)
; NAME/KEY: variation
; LOCATION: (1)...(536)
; OTHER INFORMATION: Each N is independently any nucleotide.
US-09-552-204A-11
```

```
Query Match      24.9%; Score 289.6; DB 4; Length 536;
Best Local Similarity 76.5%; Pred. No. 4,7e-65;
Matches 404; Conservative 0; Mismatches 118; Indels 6; Gaps 4;
```

```
QY 631 ATCAAGTTTGACAAGATTCTGATGAAAGAGGGTGGCCACTACATGCTTCACGGGCAAG 690
DB 1 ATCAAGTTTGACAAGATTCTGATGAAAGAGGGTGGCCACTACATGCTTCACGGGCAAG 60
QY 691 TTGCTGGGGGGGCTGGCTGGGATCTACT--ACTTCACTTACGATGACGCG--TGGCCAAAC 747
DB 61 TTGCTGGAGCGGTGGCGGGGATCTNATTAACCTTATGACATTAACGCTTGGCCAAAC 120
QY 748 AAGCACTGGCCATCGGCTGGTGACAACGGCCAGTACCGCATCCGACCTTGTGATGCC 807
DB 121 AAGCACTGTCGATCGGCTGGTGACAACGGCCAGTACCGCATCCGACCTTGTGATGCC 180
QY 808 AACACCGGCAACCAAGATGTCGCTTCAAGGCTTCAACATCTGCTGTCAAGCAGGTTGAC 867
DB 181 AACACCGGCAACCAAGATGTCGCTTCAAGGCTTCAACATCTGCTGTCAAGCAGGTTGAT 240
QY 868 GAAGTTGGTGGATGATCTTCTACTAGAGACAGAAAGGGGCTTCTATGACCTTACTGG 927
DB 241 GAAGTTGGTGGATGATCTTCTACTAGAGACAGAAAGGGGCTTCTATGACCTTACTGG 300
QY 928 ACAGACAGGCTCTTTACGGGCTTCTATCTATGATGCGACCAAGATGACCCCAAGAGTA 987
DB 301 ACAGACAGGCTCTTTACGGGCTTCTATCTATGATGCGACCAAGATGACCCCAAGAGTA 360
QY 988 TAGACATGCCACGG--CGGTCTTCACAGGACGAGGAACAAGCTTGTGACTTGGCTTACAGA 1046
DB 361 TAGACATGCCACGG--CGGTCTTCACAGGACGAGGAACAAGCTTGTGACTTGGCTTACAGA 418
QY 1047 GCAAGACCCCAACACTGATGAGCTGGGGGCTGGAGTGAAGTGAAGCGGTTTACGCTCAGG 1106
DB 419 GCAAGATCCCTCGACCTGGGCTGTGTGACATGACATCTTGGGATCTTTATTTCCAGGCAAG 478
QY 1107 CTCACCTCCCTCCGCTCTTTTTCCTTCAATTAATCCAAACCTTT 1154
DB 479 GCCTCCTTATTTGCTGCTTAAAAAAGAAATCATTAATCCAAACCTTT 526
```

```
RESULT 9
US-09-227-357-51
```

```
; Sequence 51, Application US/09227357
```

```
; Patent No. 6342581
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Fischer et al.
```

```
; TITLE OF INVENTION: 123 Human Secreted Proteins
```

```
; FILE REFERENCE: P2010P1
```

```
; CURRENT APPLICATION NUMBER: US/09/227,357
```

```
; CURRENT FILING DATE: 1999-01-08
```

```
; EARLIER APPLICATION NUMBER: PCT/US98/13684
```

```
; EARLIER FILING DATE: 1998-07-07
```

```
; EARLIER APPLICATION NUMBER: 60/051,926
```

```
; EARLIER FILING DATE: 1997-07-08
```

```
; EARLIER APPLICATION NUMBER: 60/052,793
```

```
; EARLIER FILING DATE: 1997-07-08
```

```
; EARLIER APPLICATION NUMBER: 60/051,925
```

```
; EARLIER FILING DATE: 1997-07-08
```

```
; EARLIER APPLICATION NUMBER: 60/051,929
```

```
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (485)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (486)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
```

Qy	216	GAAGAGGCTCCCTCAAC	GTGCTGCAAGCTGTGCTGGCCCCCAAGGGCCCACTCCGGCCCCC	275
Db	184	GAACTACTCCCCCAAG	TATCTGAGCAATTCCTGGCTTGGCTGAACTCCAGGGCCCCC	243
Qy	276	AGAGGCCCAAGGGCC	CTCAGCAATGATGGAGCAATGGGCTTTCCTGGCAAGAAGCGCCA	335
Db	244	TGGAGCAAAATGGTTCC	CTCGGGCCCCCATGTGTCGATCGGCTTCCAGGAAGAAGATGGTAG	303
Qy	336	AGATGSAACGA	CGGCGGACCGGGGGGACAGCGGAGAGGAAGGTCCACTGGGCCGACAGG	395
Db	304	AGAGCGCAGAAAGAG	AGAAAGGTGAAGAAAGGAACTCAGGTTTGAAGAGGTAAACCTGG	363
Qy	396	TAAACGGGGAAAGC	ACAGACCAAAAGGGCAAAAGCCGGGCGCAATTGGGCGGGCTGGCCCCG	455
Db	364	ACCGCTAGTGTTC	CGCGGTGAGAAAGGGGACCAAGAGAGACTGGGAGAAAGAACCCAT	423
Qy	456	TGGCCCAAGGGGGT	CAACGGTACCCCGGGAAGCATGGCAACCAAGGCMAAGAGGGGCC	515
Db	424	AGGACAGAGGGAG	AGAAAGAGATGATGTCMAATTGGTCTCTGGACCAAGAGGAGA	483
Qy	516	CMAAGGCAAGAAAG	GGAGGCCAGGCTCTCCAGGCCCTTCGACGCTGTGCGATGGCCATAC	575
Db	484	CMAATGATATCT	TTGGGACCCGGGGCTGCTCGAGTTTGAGATGTGAAAGCATGTGCT	543
Qy	576	CAAGTCAGCTTCT	CGGTGGCAGTGAACAAGAGCTACCAAGGGAGCGGCTGGCCCATCA	635
Db	544	CAAAATCGGCTT	TCTGTGGCATCAACAAGCTACCAAGAAANAAGCTACCTATTTAT	603
Qy	636	GTTTGACAAGATTC	T-GATGAACGAGGGTGGCCACTAACATGCTTCCA-GCGGCAAGTTC	693
Db	604	ATTTAAAGAGT	CCTCCTTCCACAGGGGAAAGCACTAACCCCTGCCACAGGGAAAGTTC	663
Qy	694	GTCTCGGCGGTGC	-TGGGATCTACTACTTCACTTACGACATCAGCGCTGGCCACAAGCA	752
Db	664	ATCTGTCTTCC	AGGGGATCTATTACTTTTCTTATGATATCAATTTGGCTAAATAAGCA	723
Qy	753	CTTGGCCATCGG	CGCTGTGCAACAACGGGCAAGTACCGCATCCGGGACCTTGAATGCCAACAC	812
Db	724	TCTGGCAATCGG	ATCGTGTACAAATGGGCAATACCGGATTAAGACCTTTCGAGGCCAACGC	783
Qy	813	CGGCAACCA	CATGTGGGCTCAGGCTCAACCATCTGGCTCTCAAGCAGGGTGAAGAAGT	872
Db	784	AGGAAACCATAT	GTGGCTTGGGGTCCACATCATCTATCTGCAACCAAGAAATGAAGT	843
Qy	873	TTGGCTGCAATCT	TCTTACAGACAGAAAGGCTCTTCTTATGAACCTTACTGTGACAGA	932
Db	844	CTGGCTGGAGATTT	CTTTCACAGACCAAGATGGGCTCTTCTTCAGACCCAGGTTGGGACAGA	903
Qy	933	CAGCCTCTTTA	CGGGCTCTCTAATCTATAGCCGACCAAGATGACC	976

Query Match	8.7%	Score 101.2;	DB 3	Length 729;
Best Local Similarity	32.6%	Pred. No. 1,3e-16;		
Matches 219, Conservative	88;	Mismatches 350;	Indels 15;	Gaps 3

313 GGCCTTCTCGCAAGAAGCGCCAAAGATGACACGACGGCGACCGGGGGGACACGCGAAG 372  
 Db 43 GGNMNCNCNCTNYTNGAYGAYAAAYABATHCCNMSNTYNTYTCGNGACACCNCGNANTN 102  
 Qy 373 GAAGTCCACCTGGCCCGGACAGGTAAACGGGGGAAAGCCAGGACCAAAAGGGCAAAAGCCGG 432  
 Db 103 CCGNAGAACNCGNGCAYCAYGAGMWSNARGGNTTNCNGMNGNGATGANNNGAYGGN 162  
 Qy 433 GCCATTGGCGGGGCTGGCCCCCGTGGCCCCCAAGGGGGGTCAAACGTATCCCCCGGGAAGCAT 492  
 Db 163 MGNATAGGNGNCNCGNGGNCNCGNGNABABAGGNGARGGNGGNGMNCNGNANTNCCN 222  
 Qy 493 GGCACACCAAGGCAAGAAAGGGGCCCAAGGGCAAGAAAGGGAGACGAGCTCCACAGCCCC 552  
 Db 223 GGNCCMNGNGNGAYCCNGGNCMMNGMNGNARGGNGCNCNGCNGNCCNAACNAGNCCN 282  
 Qy 553 TGCACTGTGGCAGTGGCCATPACCAAGTCAGCTTCTCGGTGGCAGTACCAAGAG---- 608  
 Db 283 GCGNGNGARTYMSNGTNCNCCMMGMSNCGNTTYMSNGCNAABMGMSNABRMSNGN 342  
 Qy 609 --CTACCCACGGAGCGGCTGCCCATCAAGTTTGACAAATTCTGATGAACGAGGGTGGC 666  
 Db 343 GTNCCNCCNCCNMSNGAYGCNCCNYTNCNTTYGAYMNGNTYNTGTAAAGARCARGNN 402  
 Qy 667 CACTACATATGCTTCAGCGGCAAGTGTGTCGCGCGCGCTCGGGATCTACTACTTCAAC 726  
 Db 403 CATTAATGAYGNCNGTNAACNGNAATTTACNTGYCARGTNCCNGSNGTATATATATTTGCGN 462  
 Qy 727 TAGACATCAAGCTGGCCCAACGAACCAAGCACTGGCATGGCCTGTGTACACAACGCCCATAC 786  
 Db 463 GTNCAAGCNAACNGTNTATMGNGCMWSNTYNTCAATTYGAYTNGTAAABAAAGNGABRWSN 522  
 Qy 787 CGCATCCGAGACCTTGATGCCAACACCGG-----CAACACGATGTGTGGCCTCAGGCTCC 840  
 Db 523 ATTHCCNMSNTTYTTCATTTTGTGGNGNTGGCCNAARCCNGCNCMSNTYTMWSNNGGN 582  
 Qy 841 ACCATCTGTGCTTCAAGCAGGGGTGACGAAATTTGTGTCAGATCTTCTACTCAGAGCAG 900  
 Db 583 GCNATGTGNNGNTNGARCCNGARGAYCARGTNTGGGNTNGARGTNGGTGNGAYATAY 642  
 Qy 901 AACGGGCTTTCTATGACCTTACTGACACAGACAGCTTTTACGGGCTTCTATCTAT 960

Db 643 ATGGG--NATHYACNMSNATHAARACGAYWSNACNTTYSNGNTTYTNGTATAY 639  
QY 961 GCCGACGAGAT 972  
Db 700 WSNAGATGGCAY 711

## RESULT 11

US-09-686-838B-10  
; Sequence 10, Application US/09686838B  
; Patent No. 6482612  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Humes, Jacqueline M.  
; TITLE OF INVENTION: Adipocyte-Specific Protein Homologs  
; FILE REFERENCE: 97-49D1  
; CURRENT APPLICATION NUMBER: US/09/686,838B  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/140,804  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: US 60/056,983  
; PRIOR FILING DATE: 1997-08-26  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 729  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Degenerate nucleotide sequence encoding the zslg39  
; OTHER INFORMATION: polypeptide of SEQ ID NO:2.  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (1)..(729)  
; OTHER INFORMATION: n is any nucleotide  
US-09-686-838B-10

Query Match 8.7%; Score 101.2; DB 4; Length 729;

Best Local Similarity 32.6%; Pred. No. 1.3e-16;  
Matches 219; Conservative 88; Mismatches 350; Indels 15; Gaps 3;

QY 313 GGCCTTCCTGGCAAAAGCGCCAAAGATGACACGACGCGCAACGGGGGGAACGCGAGAG 372  
Db 43 GGGWMSNCNCNTYNGAYGAAYAAATGHCNWSNTYNTGTCGNGNCAYCCGNGNYTN 102  
QY 373 GAAGGTCCACTGGCGGACAGGTAAACGGGGGAAAGCAGACCAAGGCGAAAGCCGGG 432  
Db 103 CCNGNACNCCNGNCAYCAAYGGWMSNCAAGNYTNCCGNGNNGAYGGNNGAYGN 162  
QY 433 GCCATTGGGGGGGCTGGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACCCCCGGGAAGCAT 492  
Db 163 MNGAYGNGNCCNGNNGNCCNGNNGNBARAGNBARAGNNGNNGNCCNGNNGNYTNCCN 222  
QY 493 GGCACACACGAGCAAGAAAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGCTTCCAGGCC 552  
Db 223 GGNCCMNGNNGNAYCCNGNCCMNGNNGNNGNNGNNGNCCNGNCCNACNGNCCN 282  
QY 553 TGGAGCTGTGGCAGTGGCCATCAAGTCAAGCTTTCGGTGGCAGTGAACCAAGAG-- 608  
Db 283 GCGNNGAGTGTYSNNGTNCNCCNMGWSNCGNTTYSNCCNAARMGWSNBARWSN 342  
QY 609 --CTACCCACGGAGCGGCTGCCATCAATTTGACAAGATTCTGATGAAGAGGGTGGC 666  
Db 343 GTNCCNCCNCCNMSNNGAYGNCNCTYNTCTTGTGAYGNGTNTYNGNAYGARGARGN 402  
QY 667 CACTACATGCTTCCAGCGGCAAGTTGCTGCGGCGTGCCTGGAGTCTACTACCTTACC 726  
Db 403 CATTAYGAYGNCNTNACNNGNAAATTTAATGTCARGTNCNGNNGTNTATYATTTGNC 462  
QY 727 TAGGACATCAAGCTGGCGCAACAGCAAGCTGGCCATGGCCTGTGTGCAACAGCGCATAC 786  
Db 463 GTNCAVGCNACNGTNTATYMGNCNWSNYTNCARTTYGAYTNTNARAAYGGAGARWSN 522

QY 787 CGCATCGGACCTTGTATGCAACACCGG-----CAACCAAGATGTGGCTCAGCTCC 840  
Db 523 ATHGCMWSNTTYTTCARTTYYTGGNGGNTGGCCNAARCCNCCNWSNTYTNMSNGNNGN 582  
QY 841 ACCATCTGGCTCTCAAGCAGGCTGACGAAATTGGCTGCAATCTTCTACTAGAGCAG 900  
Db 583 GCNATGTGNGNNTNBARCCNGARGAYCARGTGTGGTNCARGTNGNNGNNGAYTAY 642  
QY 901 AACGGGCTCTTATGACCTTACTGACAGACAGCGCTCTTACGGGCTTCTATCTAT 960  
Db 643 ATGGG--NATHYACNMSNATHAARACGAYWSNACNTTYSNGNTTYTNGTATAY 639  
QY 961 GCCGACGAGAT 972  
Db 700 WSNAGATGGCAY 711

## RESULT 12

US-09-140-804-9  
; Sequence 9, Application US/09140804  
; Patent No. 6197930  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Humes, Jacqueline M.  
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS  
; FILE REFERENCE: 97-49  
; CURRENT APPLICATION NUMBER: US/09/140,804  
; CURRENT FILING DATE: 1998-08-26  
; EARLIER APPLICATION NUMBER: 60/056,983  
; EARLIER FILING DATE: 1997-08-26  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 4517  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-140-804-9

Query Match 8.3%; Score 96.8; DB 3; Length 4517;

Best Local Similarity 49.7%; Pred. No. 3.5e-15;  
Matches 319; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

QY 328 GACGGCCAAAGTGAACAGACGCGACCGGGGGGACAGCGGAGAGGAGGTCACCTGCG 387  
Db 129 GCTTGACAGAGTTGATGAGGGGCAATCCAGGGCATCCGGGCCATATAGGGGCCAGAGC 188  
QY 388 CGGACAGGTAAACGGGGAAAGCCAGACCAAGAGGCGAAAGCCGGGCAATGGGGCGGCT 447  
Db 189 CGTGAATGGCAGAGATGGCACCCCTGTGAGAAAGGGTGAAGAAAGAGATCCAGGTCTTATT 248  
QY 448 GGGCCCCGTGGCCCCCAAGGGGGGTCAACGGTACCCCGGGAAGCATGGCAACAGGCAAG 507  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGAGATACCCGGGGCTGAAGGTCCCGAGGCTTT 308  
QY 508 AAGGGGCCCAAGGGCAAGAAAGGAGGACAGGCTTCCAGGCCCTGCAAGTGTGGCAGT 567  
Db 309 CCGGAATCAACAGGACGAGAAAGGAAAGCTGGAAGAGTGCCTATATATACCGC----- 362  
QY 568 GGCATATCAAGTCACTTCTCGGTGSCAGTGAACCAAGCTAACCCACGGAGCGGCTG 627  
Db 363 -----TCAGCATTCAGTGTGGATTTGGAATCTTACCTTATCTTATCCCAATG 410  
QY 628 CCGATCAAGTTGACAAAGATTCTGATGAAGAGGGTGGCCACTTACAGCTTCCAGCGG 687  
Db 411 CCGATTGCTTTACCAAGATCTTCTTACATCAGCAAAACCACTATATAGGCTTCACTGGT 470  
QY 688 AAGTGTGCTGCGGCGTGGCTGGAGTCTACTTCACTTACAGACATCAAGCTGGCGAAC 747  
Db 471 AATTTCACTGCAAGCAATTCCTGGGCTGTACTTCTTACCTTACCAATCAAGTCTATATG 510  
QY 748 AAGCACTGGCCATCGGCTGTGTGCAACAGCGCATGACGATCCGAGCTTGTATGCC 807

Db 531 AAGATGTGAAGGTGACGCTCTTCAAGAGAAGCAAGGCTATGCTTCACTATGATCAG 590  
Qy 808 AACACGGGCAACCA---CGATGTGGCTTCAGGCTCCACCATCTGAGTCTCAAGAGGCT 864  
Db 591 TACAGAGAAATATATGTGACACAGGCTCCGAGCTGTGTCTCTGATCTGAGAGTGGGC 650  
Qy 865 GACGAATTTGGCTGAGATCTTCTACTGAGAGAGAAGAGGCTCTTCTATGACCTTAC 924  
Db 651 GACCAAGTGTGCTCCAGAGGTATGAGGAGAGAGAGGCTATGAGCTCTATGCTGATAT 710  
Qy 925 TGGACAGACAGCTCTTATAGGAGCTTCTATCTATGCTGAGC 966  
Db 711 GACCAATGACTCCACCTTCAAGAGCTTCTCTCTACATGATC 752

RESULT 13  
US-09-686-838B-9  
; Sequence 9, Application US/09686838B  
; Patent No. 6482612  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Humes, Jacqueline M.  
; TITLE OF INVENTION: Adipocyte-Specific Protein Homologs  
; FILE REFERENCE: 97-49D1  
; CURRENT APPLICATION NUMBER: US/09/686,838B  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/140,804  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: US 60/056,983  
; PRIOR FILING DATE: 1997-08-26  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PaeSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 4517  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-686-838B-9

Query Match 8.3%; Score 96.8; DB 4; Length 4517;  
Best Local Similarity 49.7%; Pred. No. 3.5e-15;  
Matches 319; Conservative 0; Mismatches 302; Indels 21; Gaps 2;  
Qy 328 GACGGCCAGATGAGACAGACGAGCCGGGGGAGACGGAGAGAGAGGATCCAGCTGGC 387  
Db 129 GCTTGCACAGGTGATGAGGGGAGATCCAGGGCATCCGGCCATATAGGGGCCCAAGC 188  
Qy 388 CGGACAGATTAACGGGGGAAAGCAGACCAAGGGGCAAGCCGGGGCCATTTGGGGGCT 447  
Db 189 CGTATGTGAGAGATGAGCACTCTGTGTAGAGAGGTGAGAGAGATCCAGGTCTTAT 248  
Qy 448 GGGCCCCGTGGCCCAAGGGGGTCAACGTTACCCCGGGAAGCATGACACCAAGGCAAG 507  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGGAGTACCGGGGTGAGAGGTCCCGAGGCTTT 308  
Qy 508 AAGGGGCCCCAAGGAGAGAGAGAGAGAGGAGGCTCCAGGCCCTGAGCACTGTGAGT 567  
Db 309 CCGGGAATCCAAAGGCAAGAAAGAACTGAGAGAGGTCTTATGTATACCG- 362  
Qy 568 GGGCATACCAAGTCAAGTCTTCTCGTGGAGTGAACCAAGATCAACCAAGGAGGAGCTG 627  
Db 363 -----TCAGATTCAAGTGTGAGATTGAGATCTTACCTTATCCCAATG 410  
Qy 628 CCCATCAAGTTTGAACAAGATTCTGATGAACGAGGAGGAGCACTAAGTCTTCCAGGGC 687  
Db 411 CCATTCGCTTTTCAAGATCTTCTACATCAAGCAAAACATATGATGCTCACTGAT 470  
Qy 688 AAGTGTCTGCGGGGCTGAGATCTTACTTACCTTACCAATCAAGCTGAGGCAAC 747  
Db 471 AATTTCACTGCAACATTTCTGCGGTGTACTTACTTACCAATCAAGTCTTATATG 530  
Qy 748 AAGCACTGAGCACTGAGCTGTGACCAAGGAGAGATCAAGGAGCTGAGGAGCTTATGATGC 807  
Db 531 AAGATGTGAAGGTGAGCTCTTCAAGAGAGAGAGAGGCTATGCTTCACTATGATCAG 590

Qy 808 AACACGGGCAACCA---CGATGTGGCTTCAGGCTCCACCATCTGAGTCTCAAGAGGCT 864  
Db 591 TACAGAGAAATATATGTGACACAGGCTCCGAGCTGTGTCTCTGATCTGAGAGTGGGC 650  
Qy 865 GACGAATTTGGCTGAGATCTTCTACTGAGAGAGAAGAGGCTCTTCTATGACCTTAC 924  
Db 651 GACCAAGTGTGCTCCAGAGGTATGAGGAGAGAGAGGCTATGAGCTCTATGCTGATAT 710  
Qy 925 TGGACAGACAGCTCTTATAGGAGCTTCTATCTATGCTGAGC 966  
Db 711 GACCAATGACTCCACCTTCAAGAGCTTCTCTCTACATGATC 752

RESULT 14  
US-09-776-976-5  
; Sequence 5, Application US/09776976  
; Patent No. 6566332  
; GENERAL INFORMATION:  
; APPLICANT: Fruebis, Joachim  
; APPLICANT: Erickson, Mary Ruth  
; APPLICANT: Yen, Frances  
; APPLICANT: Bihain, Bernard  
; TITLE OF INVENTION: OB3 Globular Head and Uses Thereof for Decreasing Body Mass  
; FILE REFERENCE: 76.USA.REG  
; CURRENT APPLICATION NUMBER: US/09/776,976  
; CURRENT FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 09/758,055  
; PRIOR FILING DATE: 2001-01-10  
; PRIOR APPLICATION NUMBER: US 60/176,228  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: US 60/198,087  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: US 60/299,881  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 4517  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-776-976-5

Query Match 8.3%; Score 96.8; DB 4; Length 4517;  
Best Local Similarity 49.7%; Pred. No. 3.5e-15;  
Matches 319; Conservative 0; Mismatches 302; Indels 21; Gaps 2;  
Qy 328 GACGGCCAGATGAGACAGACGAGCCGGGGGAGACGGAGAGAGAGGATCCAGCTGGC 387  
Db 129 GCTTGCACAGGTGATGAGGGGAGATCCAGGGCATCCGGCCATATAGGGGCCCAAGC 188  
Qy 388 CGGACAGATTAACGGGGGAAAGCAGACCAAGGGGCAAGCCGGGGCCATTTGGGGGCT 447  
Db 189 CGTATGTGAGAGATGAGCACTCTGTGTAGAGAGGTGAGAGAGATCCAGGTCTTAT 248  
Qy 448 GGGCCCCGTGGCCCAAGGGGGTCAACGTTACCCCGGGAAGCATGACACCAAGGCAAG 507  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGGAGTACCGGGGTGAGAGGTCCCGAGGCTTT 308  
Qy 508 AAGGGGCCCCAAGGAGAGAGAGAGAGAGGAGGCTCCAGGCCCTGAGCACTGTGAGT 567  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGGAGTACCGGGGTGAGAGGTCCCGAGGCTTT 308  
Qy 568 GGGCATACCAAGTCAAGTCTTCTCGTGGAGTGAACCAAGATCAACCAAGGAGGAGCTG 627  
Db 309 CCGGGAATCCAAAGGCAAGAAAGAACTGAGAGAGGTCTTATGTATACCG- 362  
Qy 628 CCCATCAAGTTTGAACAAGATTCTGATGAACGAGGAGGAGCACTAAGTCTTCCAGGGC 687  
Db 411 CCATTCGCTTTTCAAGATCTTCTACATCAAGCAAAACATATGATGCTCACTGAT 470  
Qy 688 AAGTGTCTGCGGGGCTGAGATCTTACTTACCTTACCAATCAAGCTGAGGCAAC 747

```

Db      471  AATTCACTGCAACATTCCTGGCTGTACTACTTTGCTACACATCAAGTCTATATG 530
Qy      748  AAGACCTGGCAATCGGCTGTGTGCAACAGGCCAGTACCGCATCCGAGCTTTGATGCC 807
Db      531  AAGGATGTGAAGTCACTCTTTCAGAGAGACAGGCTATGCTTCCACTATAGATCAG 590
Qy      808  AACACCGGCAACCA---CGATGTGGCTCAGGCTCCACATCCTGAGCTTCAAGCAGGT 864
Db      591  TACCAAGAAAATATATGTGACACAGGCTCCGGGCTCTGTGCTCTGTGATCTGAGAGTGGC 650
Qy      865  GACGAAGTTGGCTGTGAGATCTTCTACTCAGACAGAACGGGCTTTCTATGACCTTAC 924
Db      651  GACCAAGTCTGGCTCCAGGTGTATGGGAAAGAGAGGCTATAGACTCTATGCTGATAAT 710
Qy      925  TGACAGACAGCCTTTTACGGGCTTCTATCTATGCTGAC 966
Db      711  GACAAATGACTCCACCTTCACAGGCTTCTCTTCAACATGAC 752

```

## RESULT 15

```

US-09-547-5
; Sequence 5, Application US/0909547
; Patent No. 6579852
; GENERAL INFORMATION:
; APPLICANT: Friebis, Joachim
; APPLICANT: Erickson, Mary Ruth
; APPLICANT: Yen, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: OB3 Globular Head and Uses Thereof for Decreasing Body Mass
; FILE REFERENCE: 76.US6.CIP
; CURRENT APPLICATION NUMBER: US/09/909,547
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 09/776,976
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/758,055
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: US 60/299,881
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 60/198,087
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: US 60/176,228
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 4517
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-547-5

```

```

Query Match      8.3%; Score 96.8; DB 4; Length 4517;
Best Local Similarity 49.7%; Pred. No. 3.5e-15;
Matches 319; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

```

```

Qy      328  GACGGCCCAAGATGACAGACGCGGACCGGGGGAGACAGCGAGAGAGAGTCCACTGGC 387
Db      129  GCCTGCACAGGTGTGATGGCGGATCCAGGGCATCCGGGCATATGGGGCCCCAGGC 188
Qy      388  CGGACAGGTAAACGGGGAAAGCCAGACCAAGGGCAAACCGGGGCCATTGGGCGGGCT 447
Db      189  CTGATGTGGAGAGATGGCACCCCTGTGTGAGAAAGGGTGAAGAAAGAGATCCAGGCTTAT 248
Qy      448  GGCCCCCGTGGCCCAAGGGGGTCAACGTTACCCCGGAAAGCATGGCAACACAGGCAAG 507
Db      249  GGTCTTAAGGAGACATCGTGTGAACCGAGTACCCGGGGCTGAAGTCCCGAGGCTTT 308
Qy      508  AAGGGGCCCAAGGGCAAAGAAAGGGAGCCAGGCTCCAGGCCCTGACGCTGTGGCAGT 567
Db      309  CCGGGAATCCAAGGAGAGAAAGGAAAGCTGAGAAAGTGTCTATATATACCGC----- 362
Qy      568  GGGCATACCAAGTCAAGTCTTCTGCTGTGCACTGACCAAGAGCTAACCAAGGAGCGGCTG 627
Db      363  -----TCAGCATTCAGTGTGGATGTGAAGCTTACGTTACTATCCCAACATG 410

```

```

Qy      628  CCCATCAAGTTTGACAAGATTTGATGAACGAGGTTGGCCACTATACATGCTTCCAGGGC 687
Db      411  CCAATTGCTTTACAAAGATCTTCTACATACAGCAAAACCATATGATGCTTCACTGGT 470
Qy      688  AAGTTGCTGCGGCGTGGCTGTGGAATCTACTACTTCACTACAGATCAAGCTGGCCAC 747
Db      471  AATTCACTGCAACATTCCTGGGCTGTACTACTTTGGCTTACCAATACAGTCTATATG 530
Qy      748  AACCACTGGCCATCGGCTGTGTGCAACAGGCAATCCGATCCGAGCTTGTATGCC 807
Db      531  AAGATGTGAAGTCAAGCTTCAAGGAAAGAGAGGCTATAGCTTCACTATGATCAG 590
Qy      808  AACACCGGCAACCA---CGATGTGGCTCAGGCTCCACATCCTGAGCTTCAAGCAGGT 864
Db      591  TACCAAGAAAATATATGTGACACAGGCTCCGGGCTCTGTGCTCTGTGATCTGAGAGTGGC 650
Qy      865  GACGAAGTTGGCTGTGAGATCTTCTACTCAGACAGAACGGGCTTTCTATGACCTTAC 924
Db      651  GACCAAGTCTGGCTCCAGGTGTATGGGAAAGAGAGGCTATAGACTCTATGCTGATAAT 710
Qy      925  TGACAGACAGCCTTTTACGGGCTTCTATCTATGCTGAC 966
Db      711  GACAAATGACTCCACCTTCACAGGCTTCTCTTCAACATGAC 752

```

Search completed: December 25, 2004, 06:42:11  
Job time : 128 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2004, 05:29:28 ; Search time 708 Seconds  
(Without alignments)  
9130.890 Million cell updates/sec

Title: US-10-621-787-1

Perfect score: 1161

Sequence: 1 ggaataactgctctggggcc.....aatccaactttttatca 1161

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4105333 seqs, 2784095677 residues 8210666

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :	Published Applications NA:*
1:	/cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
3:	/cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
4:	/cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
5:	/cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
6:	/cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
7:	/cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
14:	/cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
15:	/cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
16:	/cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
17:	/cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
18:	/cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
19:	/cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
20:	/cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
21:	/cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1161	100.0	1161	16	US-10-621-787-1 Sequence 1, Appl1
2	1154.2	99.4	1171	16	US-10-220-120-143 Sequence 143, Appl
3	1148.4	98.9	1178	16	US-10-162-335-29 Sequence 29, Appl
4	1120.8	96.5	1171	16	US-10-264-049-444 Sequence 44, Appl
5	1084.8	93.4	1211	15	US-10-411-120-400 Sequence 40, Appl
6	1082.8	93.3	1114	15	US-10-411-120-12 Sequence 12, Appl
7	988.4	85.1	993	15	US-10-236-055A-5 Sequence 5, Appl1
8	703.6	60.6	885	15	US-10-236-055A-7 Sequence 7, Appl1
9	702.2	60.5	1107	10	US-09-866-050A-217 Sequence 217, App
10	701.6	60.4	1107	14	US-10-152-661-217 Sequence 217, App
11	701.6	60.4	1052	10	US-09-866-050A-358 Sequence 358, App
12	701.6	60.4	1052	14	US-10-152-661-358 Sequence 358, App

## ALIGNMENTS

RESULT 1  
US-10-621-787-1  
; Sequence 1, Application US/10621787  
; Publication No. US20040024187A1  
; GENERAL INFORMATION:  
; APPLICANT: Piddington, Christopher S.  
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG  
; FILE REFERENCE: 99-08D1  
; CURRENT APPLICATION NUMBER: US/10/621,787  
; CURRENT FILING DATE: 2003-07-17  
; PRIOR APPLICATION NUMBER: US 09/552,204  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: US 60/130,207  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1161  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (133)...(987)  
US-10-621-787-1

Query Match. 100.0%; Score 1161; DB 16; Length 1161;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATACTATGCTCTGGGGCCGACGCTCTGCTCCGCTGCTGCGTGAAGACCGGGAC 60  
DB 1 GGAATACTATGCTCTGGGGCCGACGCTCTGCTCCGCTGCTGCGTGAAGACCGGGAC 60

QY 61 GCGAGAGCCCGGAGAGCTTTGCTCCGAGAGCCCTGAGAGTGGCGGAGCGCG 120  
DB 61 GCGAGAGCCCGGAGAGCTTTGCTCCGAGAGCCCTGAGAGTGGCGGAGCGCG 120  
QY 121 AGGATTAACACATGATCCCTGGGAGTCTCTGGCTGTGGCCCTGTGTCTGTAC 180  
DB 121 AGGATTAACACATGATCCCTGGGAGTCTCTGGCTGTGGCCCTGTGTCTGTAC 180  
QY 181 CCACTGCTTGGCGCTTTGCTCGAGAGGAGCTTCCGAGAAAGCTCCCTCAACTGTGTGC 240  
DB 181 CCACTGCTTGGCGCTTTGCTCGAGAGGAGCTTCCGAGAAAGCTCCCTCAACTGTGTGC 240  
QY 241 AGCTGCTGCTCCCGGAGAGCCCGGAGAGCCCGGAGAGCCCGGAGAGCCCGGAGAG 300  
DB 241 AGCTGCTGCTCCCGGAGAGCCCGGAGAGCCCGGAGAGCCCGGAGAGCCCGGAGAG 300  
QY 301 ATGGAGAGGATGGGCTTTCTGTGCAAGAGCGGCCAAGATGAGACGAGCGGAGCGGG 360  
DB 301 ATGGAGAGGATGGGCTTTCTGTGCAAGAGCGGCCAAGATGAGACGAGCGGAGCGGG 360  
QY 361 GACAGCGGAGAGGAGGCTCACTGTGCGAGAGGATGAGTAAAGCGGAGAGCGAGCAAG 420  
DB 361 GACAGCGGAGAGGAGGCTCACTGTGCGAGAGGATGAGTAAAGCGGAGAGCGAGCAAG 420  
QY 421 GCGAAGAGCCGGGAGGCTTTGGGCGGAGCGCCCGGAGAGCGGAGGCTCAAGTACC 480  
DB 421 GCGAAGAGCCGGGAGGCTTTGGGCGGAGCGCCCGGAGAGCGGAGGCTCAAGTACC 480  
QY 481 CCGGAGAGGATGGGAGAGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
DB 481 CCGGAGAGGATGGGAGAGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
QY 541 CTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
DB 541 CTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
QY 601 ACCAAGAGTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 601 ACCAAGAGTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
QY 661 GGTGGGAGTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
DB 661 GGTGGGAGTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
QY 721 TTACCTAGAGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
DB 721 TTACCTAGAGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
QY 781 CAGTACCGAGTACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
DB 781 CAGTACCGAGTACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
QY 841 ACCATCTGTGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
DB 841 ACCATCTGTGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
QY 901 AAGGAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 901 AAGGAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
QY 961 GCGGAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
DB 961 GCGGAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
QY 1021 CAAAGCTTCTGAGCTTTGAGCTTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
DB 1021 CAAAGCTTCTGAGCTTTGAGCTTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
QY 1081 GTGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
DB 1081 GTGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
QY 1141 AAATCCAAACCTTTTATCA 1161

DB 1141 AAATCCAAACCTTTTATCA 1161  
RESULT 2  
US-10-220-120-143  
Sequence 143, Application US/10220120  
Publication No. US2004048253A1  
GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: PANZER, Scott R.  
APPLICANT: SPIRO, Peter A.  
APPLICANT: BANVILLE, Steven C.  
APPLICANT: SHAH, Purvi  
APPLICANT: CHALUP, Michael S.  
APPLICANT: CHANG, Simon C.  
APPLICANT: CHEN, Alice  
APPLICANT: D'SA, Steven A.  
APPLICANT: AMSHEY, Stefan  
APPLICANT: DAHL, Christopher R.  
APPLICANT: DAM, Tam C.  
APPLICANT: DANIELS, Susan E.  
APPLICANT: DUFOUR, Gerard E.  
APPLICANT: FLORES, Vincent  
APPLICANT: FONG, Willy T.  
APPLICANT: GREENWALT, Lila B.  
APPLICANT: HITLMAN, Jennifer L.  
APPLICANT: JONES, Anissa L.  
APPLICANT: LIT, Tommy F.  
APPLICANT: ROSEBERRY, Ann M.  
APPLICANT: ROSEN, Bruce H.  
APPLICANT: RUSSO, Frank D.  
APPLICANT: STOCKREHER, Theresa K.  
APPLICANT: DAPRO, Abel  
APPLICANT: WRIGHT, Rachel J.  
APPLICANT: YAP, Pierre B.  
APPLICANT: YU, Jimmy Y.  
APPLICANT: BRADLEY, Diana L.  
APPLICANT: BRATCHER, Shawn R.  
APPLICANT: CHEN, Wensheng  
APPLICANT: COHEN, Howard J.  
APPLICANT: HODGSON, David M.  
APPLICANT: LINCOLN, Stephen E.  
APPLICANT: JACKSON, Stuart  
TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: PT-1113 PCT  
CURRENT APPLICATION NUMBER: US/10/220,120  
PRIOR FILING DATE: 2002-08-26  
PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;  
60/184,763; 60/184,771; 60/184,813; 60/184,773; 60/184,776;  
60/184,769; 60/184,768; 60/184,837; 60/184,657; 60/184,841;  
60/184,772; 60/203,785; 60/185,213; 60/204,863; 60/205,221;  
60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;  
60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;  
60/205,324; 60/205,286  
PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;  
2000-05-16; 2000-05-17; 2000-05-17; 2000-05-16; 2000-05-17;  
2000-05-17; 2000-05-17  
NUMBER OF SEQ ID NOS: 422  
SOFTWARE: PERL Program  
SEQ ID NO 143  
LENGTH: 1171  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. US2004048253A1 LI:414307.1:2000FEB01  
NAME/KEY: unsure





LENGTH: 1178  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (4) .. (994)  
US-10-162-335-29

Query Match 98.9%; Score 1148.4; DB 16; Length 1178;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1160; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 GGAATCTATGCTGGGGCGAGCTGTGCGCGGCTGCTCGGTGAGAAAGCCGAGAC 60
Db 7 GGAATCTATGCTGGGGCGAGCTGTGCGCGGCTGCTCGGTGAGAAAGCCGAGAC 66
QY 61 GCGGAGCCCCCGGAGAGCTTTCTTTGCTCGGAGCGCCCTGGAAGCTGGCGGCG 120
Db 67 GCGGAGCCCCCGGAGAGCTTTCTTTGCTCGGAGCGCCCTGGAAGCTGGCGGCG 126
QY 121 AGGGTAACCAACATGATCCCTGGGTGCTGCTGGCCCTGTCCTCCCTGTCGTGAC 180
Db 127 AGGGTAACCAACATGATCCCTGGGTGCTGCTGGCCCTGTCCTCCCTGTCGTGAC 186
QY 181 CCACTGCTGGCGCTTTGCTCGGAGGAACTTCCGAAAGGCTCCCTCAACTGTGTGC 240
Db 187 CCACTGCTGGCGCTTTGCTCGGAGGAACTTCCGAAAGGCTCCCTCAACTGTGTGC 246
QY 241 AGCCTGCTGGCGCCCGAGGCGCCACCCGCGCCCGGAGAGCCCGGAGGCGCTCA 300
Db 247 AGCCTGCTGGCGCCCGAGGCGCCACCCGCGCCCGGAGAGCCCGGAGGCGCTCA 306
QY 301 ATGGGAGCAATGGGCTTTCTGCGAAAGAGCGCCCAATGAGACGAGCGCGAGG 360
Db 307 ATGGGAGCAATGGGCTTTCTGCGAAAGAGCGCCCAATGAGACGAGCGCGAGG 366
QY 361 GACAGCGAGAGAGAGGTCCACTGTGCGCGGACAGGTAAACGGGGAAAGCCAGCA 420
Db 367 GACAGCGAGAGAGAGGTCCACTGTGCGCGGACAGGTAAACGGGGAAAGCCAGCA 426
QY 421 GCGAAAGCCGGGCGCATTTGGGCGGGCTGGCGCCCGGTGGCCCAAGGGGTTCA 480
Db 427 GCGAAAGCCGGGCGCATTTGGGCGGGCTGGCGCCCGGTGGCCCAAGGGGTTCA 486
QY 481 CCGGGAGAGATGGCAACAGCGCAAGAGAGGCGCCCAAGGGGAAAGGGGAGCC 540
Db 487 CCGGGAGAGATGGCAACAGCGCAAGAGAGGCGCCCAAGGGGAAAGGGGAGCC 546
QY 541 CTCCTAGGCGCTGCAAGCTGTGCAAGTGCATTAACAGTCAAGCTTCTGCTGCA 600
Db 547 CTCCTAGGCGCTGCAAGCTGTGCAAGTGCATTAACAGTCAAGCTTCTGCTGCA 606
QY 601 ACCAAGAGTACCAACGGAGCGGCTGCCATCAAGTTTGAACAAGATTCTGATGA 660
Db 607 ACCAAGAGTACCAACGGAGCGGCTGCCATCAAGTTTGAACAAGATTCTGATGA 666
QY 661 GGTGGCCACTCAATGCTTCCAGGCGCAAGTTGCTGCGCGCGGCTGGGATTTAC 720
Db 667 GGTGGCCACTCAATGCTTCCAGGCGCAAGTTGCTGCGCGCGGCTGGGATTTAC 726
QY 721 TTCACTTACGACATCAAGCTGTGCAACAGCACTGTGCACTGTGTGCAACAGG 780
Db 727 TTCACTTACGACATCAAGCTGTGCAACAGCACTGTGCACTGTGTGCAACAGG 786
QY 781 CAGTACCGCATTCGGGACCTTTGATGCAACACCGGCAACCAAGATGTGGCTCA 840
Db 787 CAGTACCGCATTCGGGACCTTTGATGCAACACCGGCAACCAAGATGTGGCTCA 846
QY 841 ACCATCTGCTCTTCAAGGAGGTAAGAACTTTGGTGAAGATCTTCTACTGAGAC 900
Db 847 ACCATCTGCTCTTCAAGGAGGTAAGAACTTTGGTGAAGATCTTCTACTGAGAC 906
QY 901 AAGGGCTCTTATGACCTTACTGAGACAGAGCTTTTACGGGCTTCTAATCTAT 960
```

```
Db 907 AAGGGCTCTTATGACCTTACTGAGACAGAGCTTTTACGGGCTTCTAATCTAT 966
QY 961 GCCGACCAAGATACCCCAACGAGGTATGACATGCAAGCGGCTCTCCAGGAGGAA 1020
Db 967 GCCGACCAAGATACCCCAACGAGGTATGACATGCAAGCGGCTCTCCAGGAGGAA 1026
QY 1021 CAAGCTTCTGGAATTGGGGCTTACAGAGCAAGACCCCAACTGTAGGCTGGGGTGG 1080
Db 1027 CAAGCTTCTGGAATTGGGGCTTACAGAGCAAGACCCCAACTGTAGGCTGGGGTGG 1086
QY 1081 GTCAGTGAAGCGTTTACGCTTACAGGCTCACTCTCCGCTC-TTTTTTCCTTTGAT 1139
Db 1087 GTCAGTGAAGCGTTTACGCTTACAGGCTCACTCTCCGCTCCTTTTTCCTTTGAT 1146
QY 1140 TAAATCCAAACCTTTTATTTCA 1161
Db 1147 TAAATCCAAACCTTTTATTTCA 1168
```

## RESULT 4

US-10-264-049-444  
Sequence 444, Application US/10264049  
Publication No. US2004000579A1  
GENERAL INFORMATION:  
APPLICANT: Birs et al.  
TITLE OR INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: P133P1  
CURRENT APPLICATION NUMBER: US/10/264, 049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: PatentIn Ver. 3.1  
SEQ ID NO 444  
LENGTH: 1171  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-264-049-444

Query Match 96.5%; Score 1120.8; DB 16; Length 1171;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1137; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

```
QY 18 GCCGAGGCTGTGCTCCGGCTGTGCTCCGCTGAGAAAGCCGGGAGCGCGAGCC--CGCCGA 75
Db 1 GCCGAGGCTGTGCTCCGGCTGTGCTCCGCTGAGAAAGCCGGGAGCGCGAGCCGAGCAG 60
QY 76 GAGCTTCTTTGCTCCGGAAGCGCCCTGAGCGTGGCGGAGCGCGAGGGTTAACACCATG 135
Db 61 GAGCTTCTTTGCTCCGGAAGCGCCCTGAGCGTGGCGGAGCGCGAGGGTTAACACCATG 120
QY 136 ATCCCTGGGTGCTCTGCGCTGTGCTCCCTGCTGTGCTGACCACTGCTTGGCGCC 195
Db 121 ATCCCTGGGTGCTCTGCGCTGTGCTCCCTGCTGTGCTGACCACTGCTTGGCGCC 180
QY 196 TTTGCTGCGAGGAACTTCCGGAAGAGCTCCCTTCAACTGTGTCTGCAAGCTGCTG 255
Db 181 TTTGCTGCGAGGAACTTCCGGAAGAGCTCCCTTCAACTGTGTCTGCAAGCTGCTG 240
QY 256 CAGGGGCGACCGGGGCGCCCGAGAGCGCCGAGAGGCTCAAGAAATGATGGAGCAATGG 315
Db 241 CAGGGGCGACCGGGGCGCCCGAGAGCGCCGAGAGGCTCAAGAAATGATGGAGCAATGG 300
QY 316 TTTCTTGGCAAAAGCGGCAAGATGAGCAAGCGGCGAGCGGGGAGCAAGCGAGAGAA 375
Db 301 TTTCTTGGCAAAAGCGGCAAGATGAGCAAGCGGCGAGCGGGGAGCAAGCGAGAGAA 360
QY 376 GGTTCACCTGGCGGCAAGGTAAACGGGGAAGGCAAGAACAAAGGCAAAAGCGGGGCC 435
Db 361 GGTTCACCTGGCGGCAAGGTAAACGGGGAAGGCAAGAACAAAGGCAAAAGCGGGGCC 420
```

```

QY 436 ATTGGGGGGGGTGGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACCCCGGGAGCATGGC 495
Db 421 ATTGGGGGGGGTGGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACCCCGGGAGCATGGC 480
QY 496 ACACACGAGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGCGAGGCTCCAGAGCCCTGC 555
Db 481 ACACACGAGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGCGAGGCTCCAGAGCCCTGC 540
QY 556 AGCTGTGGCAGTGGCCATACCAATCAGCTTTCTCGGTGGCAGTACCAAGAGCTACCCA 615
Db 541 AGCTGTGGCAGTGGCCATACCAATCAGCTTTCTCGGTGGCAGTACCAAGAGCTACCCA 600
QY 616 CGGAGAGGGCTGCCCATCAAGTTTGACAAAGATCTGATGAAAGAGGGTGGCACTACAAAT 675
Db 601 CGGAGAGGGCTGCCCATCAAGTTTGACAAAGATCTGATGAAAGAGGGTGGCACTACAAAT 660
QY 676 GCTTCCAGCGGCAAGTTGCTGCGGGGGTGGCTGGGATCTACTTCACTTACCTAGCATC 735
Db 661 GCTTCCAGCGGCAAGTTGCTGCGGGGGTGGCTGGGATCTACTTCACTTACCTAGCATC 720
QY 736 ACCTGTGGCCAAAGCAAGCACTGGCCATCGGCTGTGGTGAACAAGGCCAAGTACCGCATCGG 795
Db 721 ACCTGTGGCCAAAGCAAGCACTGGCCATCGGCTGTGGTGAACAAGGCCAAGTACCGCATCGG 780
QY 796 ACCTTTGATGCAACACCGGCAACCAAGATGTGGCTTCAGGCTCCACATCTTGGCTCTC 855
Db 781 ACCTTTGATGCAACACCGGCAACCAAGATGTGGCTTCAGGCTCCACATCTTGGCTCTC 840
QY 856 AAGCAGGGGTGCAAGATTTGGCTGCAAGTCTTACTACAGAGCAAGAGGGCTCTTCTAT 915
Db 841 AAGCAGGGGTGCAAGATTTGGCTGCAAGTCTTACTACAGAGCAAGAGGGCTCTTCTAT 900
QY 916 GACCTTACTGAGACAGACAGCTCTTTTACGAGCTTCTTAATCTATGCGGACAGAGATGAC 975
Db 901 GACCTTACTGAGACAGACAGCTCTTTTACGAGCTTCTTAATCTATGCGGACAGAGATGAC 960
QY 976 CCCAAGAGATATGACATGCGACAGGGGTCTCCAGGCAAGGAAACAAGCTTCTGACATT 1035
Db 961 CCCAAGAGATATGACATGCGACAGGGGTCTCCAGGCAAGGAAACAAGCTTCTGACATT 1020
QY 1036 GGGGTTACAGAGCAAGACCCCAACTGTAGCTGGGGGGTGGGGGGTGAAGTGAAGCGGTT 1095
Db 1021 GGGGTTACAGAGCAAGACCCCAACTGTAGCTGGGGGGTGGGGGGTGAAGTGAAGCGGTT 1080
QY 1096 CTAGGCTCAGGCTCAGCTCCTCCGCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1155
Db 1081 CTAGGCTCAGGCTCAGCTCCTCCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140
QY 1156 TATTTCA 1161
Db 1141 TATTTCA 1146

```

```

RESULT 5
US-10-411-120-40/c
; Sequence 40, Application US/10411120
; Publication No. US2003024425A1
; GENERAL INFORMATION:
; APPLICANT: Blonder et al.
; TITLE OF INVENTION: ACRP-30-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P316P1
; CURRENT APPLICATION NUMBER: US/10/411,120
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/328,419
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: PCT/US02/32432
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 40
; LENGTH: 1211
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
US-10-411-120-40
Query Match 93.4%; Score 1084; DB 15; Length 1211;
Best Local Similarity 99.8%; Pred. No. 8e-310;
Matches 1084; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 76 GAGTTTCTTTGCTCCGAGACCCCTTGAAGTGGCGGACCGGAGAGGTAAACCAATG 135
Db 1111 GAGTTTCTTTGCTCCGAGACCCCTTGAAGTGGCGGACCGGAGAGGTAAACCAATG 1052
QY 136 ATCCCTGAGGAGTCTCTGAGCTGAGCCCTCCCTGAGCTGTGACCACTGCTTGGGCC 195
Db 1051 ATCCCTGAGGAGTCTCTGAGCTGAGCCCTCCCTGAGCTGTGACCACTGCTTGGGCC 992
QY 196 TTGCTGCAAGGAATTCCGAAAAGCTCCCTCAACTGTGTCTGACGCTTGGCTGCC 255
Db 991 TTGCTGCAAGGAATTCCGAAAAGCTCCCTCAACTGTGTCTGACGCTTGGCTGCC 932
QY 256 CAGGGCCCATCCGGCCCCCGGAGAGGCCCGGAGGCTTCAAGAAATGATGGAACGAATGGGC 315
Db 931 CAGGGCCCATCCGGCCCCCGGAGAGGCCCGGAGGCTTCAAGAAATGATGGAACGAATGGGC 872
QY 316 TTTCCTGGCAAAAGACGAGCAAGATGACACGACGAGACCGGGGGGACAGCGAGAGAA 375
Db 871 TTTCCTGGCAAAAGACGAGCAAGATGACACGACGAGACCGGGGGGACAGCGAGAGAA 812
QY 376 GGTTCACCTGAGCGGACAGGTAAACCGGGGAAAGCCAGAGACCAAGGCAAAAGCCGGGGCC 435
Db 811 GGTTCACCTGAGCGGACAGGTAAACCGGGGAAAGCCAGAGACCAAGGCAAAAGCCGGGGCC 752
QY 436 ATTGGGGGGGGTGGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACCCCGGGAGACATGGC 495
Db 751 ATTGGGGGGGGTGGCCCCCGTGGCCCCCAAGGGGGTCAACGGTACCCCGGGAGACATGGC 692
QY 496 ACACACGAGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGCGAGGCTCCAGAGCCCTGC 555
Db 691 ACACACGAGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGCGAGGCTCCAGAGCCCTGC 632
QY 556 AGCTGTGGCAGTGGCCATACCAATCAGCTTTCTCGGTGGCAGTGAACAAGAGTACCCA 615
Db 631 AGCTGTGGCAGTGGCCATACCAATCAGCTTTCTCGGTGGCAGTGAACAAGAGTACCCA 572
QY 616 CGGAGAGGGCTGCCCATCAAGTTTGACAAAGATCTGATGAAAGAGGGTGGCACTACAAAT 675
Db 571 CGGAGAGGGCTGCCCATCAAGTTTGACAAAGATCTGATGAAAGAGGGTGGCACTACAAAT 512
QY 676 GCTTCCAGCGGCAAGTTGCTGCGGGGGTGGCTGGGATCTACTTCACTTACCTAGCATC 735
Db 511 GCTTCCAGCGGCAAGTTGCTGCGGGGGTGGCTGGGATCTACTTCACTTACCTAGCATC 452
QY 736 ACCTGTGGCCAAAGCAAGCACTGGCCATCGGCTGTGGTGAACAAGGCCAAGTACCGCATCGG 795
Db 451 ACCTGTGGCCAAAGCAAGCACTGGCCATCGGCTGTGGTGAACAAGGCCAAGTACCGCATCGG 392
QY 796 ACCTTTGATGCAACACCGGCAACCAAGATGTGGCTTCAGGCTCCACATCTTGGCTCTC 855
Db 391 ACCTTTGATGCAACACCGGCAACCAAGATGTGGCTTCAGGCTCCACATCTTGGCTCTC 332
QY 856 AAGCAGGGGTGCAAGATTTGGCTGCAAGTCTTACTCAGAGCAAGAGGGCTTCTTCTAT 915
Db 331 AAGCAGGGGTGCAAGATTTGGCTGCAAGTCTTACTCAGAGCAAGAGGGCTTCTTCTAT 272
QY 916 GACCTTACTGAGACAGACAGCTTTTACGAGCTTCTTAATCTATGCGGACAGAGATGAC 975
Db 271 GACCTTACTGAGACAGACAGCTTTTACGAGCTTCTTAATCTATGCGGACAGAGATGAC 212
QY 976 CCCAAGAGATATGACATGCGACAGGGGTCTCCAGGCAAGGAAACAAGCTTCTGACATT 1035
Db 211 CCCAAGAGATATGACATGCGACAGGGGTCTTCCAGGCAAGGAAACAAGCTTCTGACATT 152
QY 1036 GGGGTTACAGAGCAAGACCCCAACTGTAGGCTGGGGGGTGAAGTGAAGCGGTT 1095

```

Db	151	GGGCTTACAGACAGAACCCCAACTGTAGCTGGGGGGTGGGGGGTTCAGGTAGCGGTT	92
Qy	1096	CTAGCGTAGGGCTCACTCTCCGCTCTTTTTTTCCCTTCATTTAAATCCAAACCTTT	1155
Db	91	CTAGCGTAGGGCTCACTCTCTCTCTCTTTTTTTCCCTTCATTTAAATCCAAACCTTT	32
Qy	1156	TATTTCA	1161
Db	31	TATTTCA	26

```

RESULT 6
US-10-411-120-12
: Sequence 12, Application US/10411120
: Publication No. US20030224425A1
: GENERAL INFORMATION:
: APPLICANT: Blondel et al.
: TITLE OF INVENTION: ACRP-30-like Polynucleotides, Polypeptides, and Antibodies
: FILE REFERENCE: PT316P1
: CURRENT APPLICATION NUMBER: US/10/411,120
: CURRENT FILING DATE: 2003-04-11
: PRIOR APPLICATION NUMBER: 60/328,419
: PRIOR FILING DATE: 2001-10-12
: PRIOR APPLICATION NUMBER: PCT/US02/32432
: PRIOR FILING DATE: 2002-10-11
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: PatentIn Ver. 3.2
: SEQ ID NO 12
: LENGTH: 1114
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-411-120-12

```

Query Match	93.3%	Score 1082.8	DB 15	Length 1114
Best Local Similarity	99.8%	Pred. No. 1.8e-309		
Matches 1084	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

QY	76	GAGCTTTCTTCTCCGAGACGCCCTTGAGACGTGACGGGACGCCGAGGGTTAACACACATG	135
Db	9	GAGCTTCTTGTCTCCGAGACGCCCTTGAGACGTGACGGGACGCCGAGGGTTAACACACATG	68
QY	136	ATCCCTTGAGGTCTCTCTGAGCTGTGACCTTCCCTGTGTCTGTACCACTGTCTTGAGCGCC	195
Db	69	ATCCCTTGAGGTCTCTCTGAGCTGTGACCTTCCCTGTGTCTGTACCACTGTCTTGAGCGCC	128
QY	196	TTTGTCTGCAGAGGACTTCCGAGAAAGAGTCCCTTCAACTGTGTGACGCTGTGAGCGCC	255
Db	129	TTTGTCTGCAGAGGACTTCCGAGAAAGTCTCCCTCAACTGTGTGTGACGCTGTGAGCGCC	188
QY	256	CAGGACCCACCCGCGCTCCCAAGAGCCTTCAAGGACCTTCAAGATGATGTGACGAATGTGAC	315
Db	189	CAGGACCCACCCGCGCTCCCAAGAGCCTTCAAGATGATGTGACGAATGTGAC	248
QY	316	TTTCTGTGCAGAAAGACGGCCAGATGTGACACGACGGCGACCGGGGGGAGACAGCGGAGAGGA	375
Db	249	TTTCTGTGCAGAAAGACGGCCAGATGTGACACGACGGCGACCGGGGGGAGACCGGAGAGGA	308
QY	376	GGTCCACCTTGCGCCGAGACAGGTAAACCGGGGAAAGCCAGGACCAAGAGGCAAAAGCCGGAGCC	435
Db	309	GGTCCACCTTGCGCCGAGACAGGTAAACCGGGGAAAGCCAGGACCAAGAGGCAAAAGCCGGAGCC	368
QY	436	ATTGGGGCGGGCTGTGCCCCCTGTGCCCCCAAGGGGGGTCAACGTATACCCCTGGGAAAGGATGGC	495
Db	369	ATTGGGGCGGGCTGTGCCCCCTGTGCCCCCAAGGGGGGTCAACGTATACCCCTGGGAAAGGATGGC	428
QY	496	ACAACAGGCAAGAAAGGGGCCCAAGGGGCAAGAAAGGGAGCCAGGCTCCAGAGCCCTTGC	555
Db	429	ACAACAGGCAAGAAAGGGGCCCAAGGGGCAAGAAAGGGAGCCAGGCTCCAGAGCCCTTGC	488
QY	556	AGCTGTGGCAATGTGCAATACCAAGTCACTTTCTGTGTGTGACGTGTACCAAGAGCTAACCA	615
Db	489	AGCTGTGGCAATGTGCAATACCAAGTCACTTTCTGTGTGTGACGTGTACCAAGAGCTAACCA	548

QY	CGGAGGGGGCGCCCATCAAGTTTGACAAAGATCTGATGAA	CGAGGGGGGCGCATCAAT	675	
Db	CGGAGGGGGCTGCCCATCAAGTTTGACAAAGATCTGATGAA	CGAGGGGGGCGCATCAAT	608	
QY	676	GCTTCAGCGGCAAGTTCGTCTGCGGGGTGCTTGGAATTACTACTTCACCTACGACATC	735	
Db	609	GCTTCAGCGGCAAGTTCGTCTGCGGGGTGCTTGGAATTACTACTTCACCTACGACATC	668	
QY	736	ACGCTGGCCAAACAAGCACCTGGCCATGGGCTGGTGCA	CGGCCAGTACCGCATCCGG	795
Db	669	ACGCTGGCCAAACAAGCACCTGGCCATGGGCTGGTGCA	CGGCCAGTACCGCATCCGG	728
QY	796	ACCTTGTATGCAACACCGGCAACACGATGTGACCTCAGGCTCCACCATCTGGCTTC	855	
Db	729	ACCTTGTATGCAACACCGGCAACACGATGTGACCTCAGGCTCCACCATCTGGCTTC	788	
QY	856	AAGCAGGGGTGACGAAGTTTGCTGACAGATCTTACTACAGACGAACCGGCTCTTCTAT	915	
Db	789	AAGCAGGGGTGACGAAGTTTGCTGACAGATCTTACTACAGACGAACCGGCTCTTCTAT	848	
QY	916	GACCTTACTGAGACAGACAGCTCTTTTACGGGCTTCTAATCTATGCGACACAGATGAC	975	
Db	849	GACCTTACTGAGACAGACAGCTCTTTTACGGGCTTCTAATCTATGCGACACAGATGAC	908	
QY	976	CCCAACAGAGTATGACATGCAAGGGGGTCTCCAGGACGGGAAACAAGTTTCGATCTT	1035	
Db	909	CCCAACAGAGTATGACATGCAAGGGGGTCTCCAGGACGGGAAACAAGTTTCGATCTT	968	
QY	1036	GAGCTTACAGACAGACCCCAACAATGTAGGCTGGGGGGTGGGGGGTCTGATGAGCGGTT	1095	
Db	969	GAGCTTACAGACAGACCCCAACAATGTAGGCTGGGGGGTGGGGGGTCTGATGAGCGGTT	1028	
QY	1096	CTAGCCTCAGGCTCACTCTCGGCTCTTTTTTCCCTTCATTTAAATCCAAACCTTTT	1155	
Db	1029	CTAGCCTCAGGCTCACTCTCTGCTCTTTTTTCCCTTCATTTAAATCCAAACCTTTT	1088	
QY	1156	TATTTCA	1161	
Db	1089	TATTTCA	1094	

```

RESULT 7
US-10-236-055A-5
; Sequence 5, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehرداد
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-055A-5

```

Query Match	85.1%	Score 988.4	DB 15	Length 993
Best Local Similarity	99.9%	Pred. No. 1.3e-281		
Matches 989	Conservative 0	Mismatches 1	Indels 0	Gaps 0

  

Qy	1	GGAAAACTATGCTCGG393CCGACGCTCTTGCCCGGCTGCTGCGCTGAGGAAAGCCCGGAC	60
db	4	GGAAAACTATGCTCGG393CCGACGCTCTTGCCCGGCTGCTGCGCTGAGGAAAGCCCGGAC	63

```

Qy 61 GCGAGAGCCCCCGGAGAGCTTCTTGTCTCGGAGCGCCCTTGAGAGTGGCGGAGCCGCG 120
Db 64 GCGAGAGCCCCCGGAGAGCTTCTTGTCTCGGAGCGCCCTTGAGAGTGGCGGAGCCGCG 123
Qy 121 AGGTTAACCAACATGATCCCTGGGATGCTCTGAGCTGTGCTCTCCCTGTGTCTGAC 180
Db 124 AGGTTAACCAACATGATCCCTGGGATGCTCTGAGCTGTGCTCTCCCTGTGTCTGAC 183
Qy 181 CCATGCTTGGCGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 240
Db 184 CCATGCTTGGCGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 243
Qy 241 AGCTGCTTGGCGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 300
Db 244 AGCTGCTTGGCGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 303
Qy 301 ATGGGAGCAATGGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 360
Db 304 ATGGGAGCAATGGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 363
Qy 361 GACAGCGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 420
Db 364 GACAGCGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 423
Qy 421 GGCAGAGCGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 480
Db 424 GGCAGAGCGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 483
Qy 481 CCGGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 540
Db 484 CCGGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 543
Qy 541 CTCGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 600
Db 544 CTCGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 603
Qy 601 ACCAGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 660
Db 604 ACCAGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 663
Qy 661 GGTGGCACTTACATGCTTCCAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGG 720
Db 664 GGTGGCACTTACATGCTTCCAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGG 723
Qy 721 TTCACTACGACATCAGCGTGGCCAAACAGCACTGGCCATCGGCTGTGTGCAACGGC 780
Db 724 TTCACTACGACATCAGCGTGGCCAAACAGCACTGGCCATCGGCTGTGTGCAACGGC 783
Qy 781 CAGTACCGCATCCGAGCTTGTGATGCCAACACCGGCAACAGATGTGGCTCAGGCTCC 840
Db 784 CAGTACCGCATCCGAGCTTGTGATGCCAACACCGGCAACAGATGTGGCTCAGGCTCC 843
Qy 841 ACCATCTGGCTTCAAGACGAGGAGTGAAGAGTTTGGCTGAGATCTTCTACTCAGAGAG 900
Db 844 ACCATCTGGCTTCAAGACGAGGAGTGAAGAGTTTGGCTGAGATCTTCTACTCAGAGAG 903
Qy 901 AACGGGCTTCTTATGACCTTACTGACAGACAGCTCTTCAAGGCTTCTTATCTAT 960
Db 904 AACGGGCTTCTTATGACCTTACTGACAGACAGCTCTTCAAGGCTTCTTATCTAT 963
Qy 961 GCGAGCAAGATGACCCCAACGAGATATAG 990
Db 964 GCGAGCAAGATGACCCCAACGAGATATAG 993

```

RESULT 8  
 US-10-236-055A-7  
 ; Sequence 7, Application US/10236055A  
 ; Publication No. US20030134328A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Basham, Beth E.  
 ; APPLICANT: Forsythe, Ian

```

; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-236-055A-7

Query Match      60.6%; Score 703.6; DB 15; Length 885;
Beet Local Similarity 88.0%; Pred. No. 1.8e-197;
Matches 766; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 121 AGGTTAACCAACATGATCCCTGGGATGCTCTGAGCTGTGCTCTCCCTGTGTCTGAC 180
Db 16 AAGTTGACACCAATGATCTCTGGGTACTCTTGGCTGTGTCTCTCGTGTGTCTGAC 75
Qy 181 CCATGCTTGGCGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 240
Db 76 CCATGCTTGGGTGCTTGTCTGAGAGGACTTCCGGAAGGAGGAGGCTCTCAACTGGTCTGC 135
Qy 241 AGCTGCTTGGCGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 300
Db 136 AGCTGCTTGGTCTTCCAGAGGCTTCCAGAGGCTTCCAGAGGCTTCCAGAGGCTG 195
Qy 301 ATGGGAGCAATGGGCTTGTCTGAGAGGACTTCCGGAAGGCTCCCTCAACTGGTCTGC 360
Db 196 GTGGAGAGAGTGGTTCCTCTGGGAAAGCGGCCAAGATGGCCAGAGCGAGGCGG 255
Qy 361 GACAGCGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 420
Db 256 GACAGTGAAGAGAGGAGTCCACTGCGAGGACAGGCAACCTGGAAACAAAGCAAG 315
Qy 421 GGCAGAGCGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 480
Db 316 GGCAGAGCTGGAGCATTTGGAGAGTGGCCCTCTGAGAGCTTCAAGGAGTGAAGTACC 375
Qy 481 CCGGAGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 540
Db 376 CCGGAGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 435
Qy 541 CTCGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 600
Db 436 CTCGAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 495
Qy 601 ACCAGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 660
Db 496 ACCAGAGGAGTCCACTGCGCGGACAGGTAAACGGGAAAGCCAGAACAAAG 555
Qy 661 GGTGGCACTTACATGCTTCCAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGG 720
Db 556 GGTGGCACTTACATGCTTCCAGAGGAGGAGTCCACTGCGCGGACAGGTAAACGGG 615
Qy 721 TTCACTACGACATCAGCGTGGCCAAACAGCACTGGCCATCGGCTGTGTGCAACGGC 780
Db 616 TTCACTATGACATTAAGCTGTGGCCAAACAGCACTGGCCATCGGCTGTGTGCAACGGT 675
Qy 781 CAGTACCGCATCCGAGCTTGTGATGCCAACACCGGCAACAGATGTGGCTCAGGCTCC 840
Db 676 CAGTACCGCATTCGAGCTTGTGATGCCAACACCGGCAACAGATGTGGCTCAGGCTCC 735
Qy 841 ACCATCTGGCTTCAAGACGAGGAGTGAAGAGTTTGGCTGAGATCTTCTACTCAGAGAG 900
Db 736 ACCATCTTACTCTCAAGAGGAGTGAAGAGTTTGGCTGAGATCTTCTACTCAGAGAG 795

```

QY 901 AACGGCTCTTCTATGACCTTTATGACAGACAGCCCTTTAAGGGCTTCTATCAT 960  
DB 796 AATGGCTCTTCTATGACACCTTTATGACAGACAGCCCTTTAAGGGCTTCTATCAT 855  
QY 961 GCCGACCAAGATGACCCCAAGAGTATAG 990  
DB 856 GCTGACCAAGAGACCCCAAGAGTATAG 885

## RESULT 9

US-09-866-050A-217  
; Sequence 217, Application US/09866050A  
; Publication No. US20030040471A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OR INVENTION: Compositions Isolated From Skin Cells  
; TITLE OR INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c40  
; CURRENT APPLICATION NUMBER: US/09/866,050A  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 217  
; LENGTH: 1107  
; TYPE: DNA  
; ORGANISM: Rat  
US-09-866-050A-217

Query Match 60.5%; Score 702.2; DB 10; Length 1107;  
Best Local Similarity 84.7%; Pred. No. 5e-197;

Matches 800; Conservative 0; Mismatches 143; Indels 2; Gaps 1;

QY 121 AGGGTAACCAACATGATCCCTGGGTGCTCTGAGCTGTGCTCCCTGTGCTGTGAC 180  
DB 162 AAGGTGACCAACATGATCTCTGATGCTCTTGGCTGTGCTCTGCTGTGCTGTGAC 221  
QY 181 CCAGTCTTGGCGGCTTGTCTGCGAGGACTTTCGGAAAGGCTCCCTCAACTGTCTGC 240  
DB 222 CCAATGCTTGGTGTCTTGTCTGCGAGGACTTTCGAAAGGCTGTCTCAACTGTCTGC 281  
QY 241 AGCTGCTGTGCTCCCGAGGCTCCCGAGGCTCCCGAGGCTCCCGAGGCTCCCGAGG 300  
DB 282 AGTGTCTGTGTCTCCCGAGGCTCCCGAGGCTCCCGAGGCTCCCGAGGCTCCCGAGG 341  
QY 301 ATGGGACCAATGGGCTTCTCTGCGAAAGACGCGCAAGTGAACAAGCGGCGGCGG 360  
DB 342 GTGGGAAAGATGGGTTTCTCTGTGAAGATGGCCCAAGCGGCGGCGGCGGCGGCGG 401  
QY 361 GACAGCGAGAGAAAGTCTCACTGCGCGAGCAAGTAAACCGGAGAAACCGAGCAAG 420  
DB 402 GACAGTGAAGAAAGGCTCACTGCGCGAGCAAGTAAACCGAGAAACCGAGCAAG 461  
QY 421 GGCAAAACCGGCGGCTTGTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCT 480  
DB 462 GGCAAAACCGGCGGCTTGTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCT 521  
QY 481 CCGGAGAGCATGAGCAACGAGCAAGAAAGGCGCAAGAAAGGCGAGCGGCGGCGG 540  
DB 522 CCGGAGAGCATGAGTATACCGGCGCAAGAGGACTTAAGGCGCAAGAAAGGCGAG 581  
QY 541 CTCCAGGCGGCTGAGCTGTGCGAGTGGCCATCAAGTCAAGCTTCTGCGGCGAGT 600  
DB 582 CTCCAGGCGGCTGAGCTGTGCGGCGAGTGGCCAGGCAAGTGGCGCTTCTGCGGCGGTA 641  
QY 601 ACAGAGAGCTTACCAAGGAGAGGCTGCGGCTCAAGTGTGAACAAGTGTGAACAAG 660  
DB 642 ACAGAGAGTACCAAGTGAAGAGCTGCCCATCAAGTGTGAACAAGTGTGAACAAGT 701

QY 661 GGTGGCACTTACATGCTTCCAGCGGCAAGTTCGTGCGGCGGCGGCGGCGGCGGCT 720  
DB 702 GAGAGCACTTACATGATGATCAAGTGGCAAGTTCGTGCGGCGGCGGCGGCGGCGGCT 761  
QY 721 TTCACTTACATGATCAAGTGGCGGCAAGCAAGCTTGGCGGCTTGGCGGCGGCGGCT 780  
DB 762 TTCACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 821  
QY 781 CAGTACCGCATCCGGAACCTTGTATGCAACCGGCGCAACCGATGTGGCTTCAGGCTCC 840  
DB 822 CAGTACCGCATTCGGAACCTTGTATGCAACCGGCGCAACCGATGTGGCTTCAGGCTCC 881  
QY 841 ACATCTGCTCTCAAGCAAGGCGGCAAGTTCGTGCGGCGGCGGCGGCGGCGGCGGCT 900  
DB 882 ACATCTGCTCTCAAGCAAGGCGGCGGCAAGTTCGTGCGGCGGCGGCGGCGGCGGCT 941  
QY 901 AACGGCTCTTCTATGACCTTTAAGGACAGACGCTTCTTACGGGCTTCTATCAT 960  
DB 942 AATGACTCTTCTATGACACCTTTAAGGACAGACGCTTCTTACGGGCTTCTATCAT 1001  
QY 961 GCCGACCAAGATGACCCCAAGAGTATAGACATGCCAGG--CGGTCTCCAGGCGAGG 1018  
DB 1002 GCTGATCAAGAGACCCCAAGAGTATAGCAAGAGTGGGCTTGAAGCGCTCCAGGCG 1061  
QY 1019 AACAGCTCTGACTTGGGCTTACAGAGCAAGACCCCAAGT 1063  
DB 1062 ACTAGATTCGCAAGGCTGTGATAGAGAGATCTGTAGACTG 1106

## RESULT 10

US-10-152-661-217  
; Sequence 217, Application US/10152661  
; Publication No. US20030022835A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OR INVENTION: Compositions Isolated From Skin Cells  
; TITLE OR INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c5  
; CURRENT APPLICATION NUMBER: US/10/152,661  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 217  
; LENGTH: 1107  
; TYPE: DNA  
; ORGANISM: Rat  
US-10-152-661-217

Query Match 60.5%; Score 702.2; DB 14; Length 1107;  
Best Local Similarity 84.7%; Pred. No. 5e-197;  
Matches 800; Conservative 0; Mismatches 143; Indels 2; Gaps 1;

QY 121 AGGGTAACCAACATGATCCCTGGGTGCTCTGAGCTGTGCTCCCTGTGCTGTGAC 180

Db 162 AAGGTGACCAACGATCTCTGATGCTCTTGGCTGTGCTCTCTGCTGCTGAC 221  
Qy 181 CCACTGCTTGGCGCTTGTCTGCAAGGACTTCCGAAAGGCTCCCTCAACTGCTGTC 240  
Db 222 CCAATGCTTGGTGTCTTGTCTGCAAGGACTTCCGAAAGGCTGCTCAACTGCTGTC 281  
Qy 241 AGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 300  
Db 282 AGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 341  
Qy 301 ATGGGACGAATGGCTTGTCTGCAAGGACTTCCGAAAGGCTCCCTCAACTGCTGTC 360  
Db 342 GTGGGAAAGATGGCTTGTCTGCAAGGACTTCCGAAAGGCTCCCTCAACTGCTGTC 401  
Qy 361 GACAGCGAG 420  
Db 402 GACAGTGAAG 461  
Qy 421 GCGAAG 480  
Db 462 GCGAAG 521  
Qy 481 CCGGAG 540  
Db 522 CCGGAG 581  
Qy 541 CTCCGAG 600  
Db 582 CTCCGAG 641  
Qy 601 ACCAAG 660  
Db 642 ACCAAG 701  
Qy 661 GGTGGCACTACAG 720  
Db 702 GGTGGCACTACAG 761  
Qy 721 TTACACTACAG 780  
Db 762 TTACACTACAG 821  
Qy 781 CAGTACAG 840  
Db 822 CAGTACAG 881  
Qy 841 ACCATCTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
Db 882 ACCATCTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 941  
Qy 901 AAGGAGCTTCTTATGACCTTATGACCTTATGACCTTATGACCTTATGACCTTAT 960  
Db 942 AAGGAGCTTCTTATGACCTTATGACCTTATGACCTTATGACCTTATGACCTTAT 1001  
Qy 961 GCGGACAG 1018  
Db 1002 GCGGACAG 1061  
Qy 1019 AACAAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAG 1063  
Db 1062 AACAAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAG 1106

## RESULT 11

US-09-866-050A-358  
; Sequence 358, Application US/09866050A  
; Publication No. US20030040471A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene

APPLICANT: Murison, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions Isolated From Skin Cells  
FILE REFERENCE: 11000.1011c4U  
CURRENT APPLICATION NUMBER: US/09/866,050A  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ. ID NOS: 725  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO: 358  
LENGTH: 1052  
TYPE: DNA  
ORGANISM: Rat  
US-09-866-050A-358

Query Match 60.4%; Score 701.6; DB 10; Length 1052;  
Best Local Similarity 82.9%; Pred. No. 7,4e-197;  
Matches 837; Conservative 0; Mismatches 169; Indels 4; Gaps 3;

Qy 121 AAGGTACCAACGATCTCTGATGCTCTTGGCTGTGCTCTCTGCTGCTGAC 180  
Db 34 AAGGTACCAACGATCTCTGATGCTCTTGGCTGTGCTCTCTGCTGCTGAC 93  
Qy 181 CCACTGCTTGGCGCTTGTCTGCAAGGACTTCCGAAAGGCTCCCTCAACTGCTGTC 240  
Db 94 CCAATGCTTGGTGTCTTGTCTGCAAGGACTTCCGAAAGGCTGCTCAACTGCTGTC 153  
Qy 241 AGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 300  
Db 154 AGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 212  
Qy 301 ATGGGACGAATGGCTTGTCTGCAAGGACTTCCGAAAGGCTCCCTCAACTGCTGTC 360  
Db 213 GTGGGAAAGATGGCTTGTCTGCAAGGACTTCCGAAAGGCTCCCTCAACTGCTGTC 272  
Qy 361 GACAGCGAG 420  
Db 273 GACAGTGAAG 332  
Qy 421 GCGAAG 480  
Db 333 GCGAAG 392  
Qy 481 CCGGAG 540  
Db 393 CCGGAG 452  
Qy 541 CTCCGAG 600  
Db 453 CTCCGAG 512  
Qy 601 ACCAAG 660  
Db 513 ACCAAG 572  
Qy 661 GGTGGCACTACAG 720  
Db 573 GGTGGCACTACAG 632  
Qy 721 TTACACTACAG 780  
Db 633 TTACACTACAG 692  
Qy 781 CAGTACAG 840  
Db 693 CAGTACAG 752  
Qy 841 ACCATCTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
Db 753 ACCATCTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812  
Qy 901 AAGGAGCTTCTTATGACCTTATGACCTTATGACCTTATGACCTTATGACCTTAT 960



Db	813	AATGAGCTCTTCTAGACCCCTTATTGACCCGACGCTGTTACGGGGCTTCTCATCTAC	872
Qy	961	GCCGACCGAGATGATACCCCAAGAGGTATAGACATGCCACGG-CGGTCTCTCCAGGACGGA	1013
Db	873	GCTGATCAAGAGAGACCCCAATGAGGTATTAGACCAAGCTGGGGTTGAGCGTCCAGGACGGGA	932
Qy	1020	ACAAAGCTTCTGGACTTGGGGCTTACAGACGAAACCCCAAACTGTAGGCTGGGGGGTGGGG	1077
Db	933	CTAAGATTCCCAAGGGGTGTGATGAAAGAGGATTTCTGMACTGAGGCTGGGGGACTGGCA	992
Qy	1080	GATCGAGTAGCGGTTCTAGCCTCAGGCTCAACCTCTCGGCTCTTTT	1129
Db	993	GTTCTGGGAGC--TTTATTTCACAGGCAAGCTCTCTTGAGTCTGCTTT	1040

```

RESULT 12
US-10-152-661-358
Sequence 358, Application US/10152661
Publication No. US20030022835A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
TITLE OF INVENTION: and Methods for Their Use
FILE REFERENCE: 11000.1011C5
CURRENT APPLICATION NUMBER: US/10/152,661
CURRENT FILING DATE: 2002-05-20
PRIOR APPLICATION NUMBER: 09/866,050
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: 60/221,232
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/206,650
PRIOR FILING DATE: 2000-05-24
PRIOR APPLICATION NUMBER: 09/312,283
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/NZ99/00051
PRIOR FILING DATE: 1999-04-29
PRIOR APPLICATION NUMBER: 09/188,930
PRIOR FILING DATE: 1998-11-09
PRIOR APPLICATION NUMBER: 09/069,726
PRIOR FILING DATE: 1998-04-29
NUMBER OF SEQ ID NOS: 725
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 358
LENGTH: 1052
TYPE: DNA
ORGANISM: Rat
US-10-152-661-358

```

Query Match	60.4%;	Score 701.6;	DB 14;	Length 1052;
Best Local Similarity	82.9%;	Pred. No. 7.4e-197;		
Matches 837;	Conservative 0;	Mismatches 169;	Indels 4;	Gaps 3

QY	121	GGGGTAAACACACATATATCCCGGGGTGCTCCGAGGCTGAGCCCTCCCGTACTGTAC	180
Db	34	AAGGTGACCAACATATCTCTCGAATGCTCTTGAGCTTGACCTTCGATGCTGTGAC	93
QY	181	CCACTGCTTGGCGGCTTGTCTGCGAGGACCTTCCGAAAAGGCTCCCTCAACTGTGCTGC	240
Db	94	CCAAATGCTTGGGCTTTGTCTGCGAGGACTTCGAGAAAGGAGTGTCTCAACTGGATGTC	153
QY	241	AGCTGCTGGGCCCCCAGGGGCCCACTCCGGCCCCCGAGAGCCCAAGGCTCTCAGAAATG	300
Db	154	AGTCTGCTGGTCCCAAGG--CMACTGGCTCTCAAGGAGCACAGATCTCTCAGAAATG	212
QY	301	ATGGAGCAATAGGCTTCTCTGCGAAAGACGGCCMAAGATGACAGATGGCGACCCGGGG	360
Db	213	GTGGGAAATATGGATTCTCTGTAGAGATGGCCAAAGCCGCGACGAGACGAAACGAGGG	272

OY	361	GACAGCGGAGAGGAAGGTCACCTGGCGCGGACAGGTAAACCGGGGAAAGCCAGGACCAAAG	420
Db	273	GACAGTGGAGAGAGAGGTCACCTGGCGAGGACAGGCAACCGAGGAAACAAAGACCAAAG	332
OY	421	GGCAAAACCCGAGGCGCATTTGGGCGGGCTGGGCCCCCGTGGCCCAAGGGGTCAAACGGTACC	480
Db	333	GGCAAAACTGGGGGCATTTGGGAGAGCGGGTCTTCGAGGACCAAGGGGGTCAGTGGTACC	392
OY	481	CCCGGAAAGCATGGCAACAACGAGCAGAGAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGC	540
Db	393	CCCGGAAACATGGTATAACCGGGGCAAGAAAGGAGACTTAAGGGCAAGAAAGGGGAAACTGGGG	452
OY	541	CTCCAGAGCCCTTGACCTGTGGGAGGGGCATATCAAGTCAGCTTCTCGGTGGCAGTG	600
Db	453	CTCCAGAGCCCTGTACCTGGGCGAGTACGCGAGCCAGGCAAGTGGCTTTTGGGTGGCGGTAA	512
OY	601	ACCAAGAGCTAACCCACCGGAGCGGGCTGCCCATCAAGTTTGAACAAGATTCTGATGAAGAG	660
Db	513	ACCAAGAGTTAACCCAGTGAGCGACTGGCCCATCAAGTTTGAACAAGATTCTGATGAATGAG	572
OY	661	GGTGGCACTAACATGCTTCCAGCGGCAAGTTGCTTGGCGCGTGGCTGGGATCTACTAC	720
Db	573	GGAGGCCACTAACATGATCATCAGTGGGAGAGTGGCTGTGCAAGCGTGGCCAGGGATCTATTAC	632
OY	721	TTTCACTTAACACATCAAGCTGGGCCCAACAAGCACTGGGCCATTCGGCCCTGGTGGCAACAAGGC	780
Db	633	TTTACCTTAATGACATTAAGCTGGGCCCAACAAACACTGGGCCATTCGGCTTAAGTGCACAAATGGC	692
OY	781	CAGTACCGGCATCCGGGACCTTTGATGTCGAACAACCGGCACAACAGATGTGGGCTCAGGCTCC	840
Db	693	CAGTACCGGCATTCGGGACTTTTGAACGCCAACACCGGCAACACAGACGTGGGCTCGGGCTCC	752
OY	841	ACCATCTGGCTCTCAAGCAGGGGTGACGAAAGTTTGGCTGCAATCTTCTACCTCAGACAG	900
Db	753	ACCATCTAGCTCTCAAGAGGGGTGATGAAGTCTGGTTACAGATTTTCTACTCGGAGCAG	812
OY	901	AACGGGCTTTTCTATGACCTTCTACTGGAACAGACGCTCTTTACGGGGCTCTCTAATCTAT	960
Db	813	AATGGACTCTTTTACGACCTCTTAATTTGAGCCGACGCTGTCTTACCGGGCTTCTCATCTAC	872
OY	961	GCACAGCAGATGACCCCAACGAGAGTATTACATGTGCACGG- CGGTCCTCAGGCAAGGA	1019
Db	873	GCTAATCAAGAGACCCCAATGAGGTATTAACAAGCTGGGGTTGAGGTCTCAGGCAAGGA	932
OY	1020	ACAAGCTTCTGGAATTGGGCTTACAGAGCAAGACCCCAACACTGTAGGCTGGGGGTGGG	1079
Db	933	CTAAGATTTCGCAAGGGGTGCTGATAGAGAGAGATCTCTGAATCTGAGGCTGGGAGCTGGCA	992
OY	1080	GGTGAATGAGCGGTTCTAGCTTCAGGCTCACTCCCTCCGCTCTTTT	1129
Db	993	GTTCCTTGGAGC--TTTATTTCCAGGCAAGCCTCTCTGGTGTGCTTT	1040

```

? RESULT 13
? US-09-770-906-1
?
? Sequence 1, Application US/09770906
? Patent No. US20010009905A1
?
? GENERAL INFORMATION:
?
? APPLICANT: HENSLEY, PRESTON
?
? APPLICANT: HU, ERDING
?
? APPLICANT: SMITH, RANDALL FORREST
?
? APPLICANT: ZHU, YUAN
?
? TITLE OF INVENTION: ACRP30R1: A HOMOLOG OF ACRP30 (30 KD
? TITLE OF INVENTION: ADIPOCYTE COMPLEMENT-RELATED PROTEIN)
?
? FILE REFERENCE: GP-70455-C1
?
? CURRENT APPLICATION NUMBER: US/09/770,906
?
? CURRENT FILING DATE: 2001-01-26
?
? PRIOR APPLICATION NUMBER: 60/086,562
?
? PRIOR FILING DATE: 1998-05-21
?
? PRIOR APPLICATION NUMBER: 09/162,352
?
? PRIOR FILING DATE: 1998-09-29
?
? NUMBER OF SEQ ID NOS: 4
?
? SOFTWARE: FastSeq for Windows Version 3.0

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:08:11 ; Search time 73.0538 Seconds  
(without alignments)  
375.364 Million cell updates/sec

Title: US-10-621-787-2

Perfect score: 1582

Sequence: 1 MIRMVLLACALPCADPLG.....TDSLFTGFLIYADDDNEV 285

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: PIR.79:\*  
2: p1r1:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	516.5	32.6	680	1 CGHUIB	collagen alpha 1(X
2	511	32.3	680	2 S31216	collagen alpha 1(X
3	502.5	31.8	674	2 S13301	collagen alpha 1(X
4	502	31.7	674	2 S23297	collagen alpha 1(X
5	443.5	28.0	635	2 A57131	collagen alpha 2(V
6	443.5	28.0	743	1 S23779	collagen alpha 1(V
7	439	27.7	744	1 A34246	collagen alpha 1(V
8	435	27.5	744	1 S23298	collagen alpha 1(V
9	431.5	27.3	744	2 S15435	gelatin-binding 28
10	420	26.5	744	2 A55797	collagen alpha 1(V
11	398.5	25.2	423	2 T14782	collagen precursor
12	393	24.8	219	2 S29328	hypothetical prote
13	376.5	23.8	253	1 S19018	complement subcomp
14	369	23.3	253	2 S19018	complement subcomp
15	369	23.3	253	2 S19018	complement subcomp
16	366	23.3	253	2 S19018	complement subcomp
17	366	23.3	253	2 S19018	complement subcomp
18	357	22.6	253	2 S49158	complement subcomp
19	348.5	22.0	245	1 S49158	complement subcomp
20	311.5	19.7	1464	1 CGHUIB	collagen alpha 1(I
21	310	19.6	1453	2 S21626	collagen alpha 1(I
22	309	19.5	671	1 CGRTIS	collagen alpha 1(I
23	306	19.3	215	2 B48150	collagen alpha 1(I
24	306	19.3	1042	1 CGCHIS	collagen alpha 1(I
25	303.5	19.2	779	1 CGBOIS	collagen alpha 1(I
26	300.5	19.0	299	2 T24833	hypothetical prote
27	296.5	18.7	266	2 T22706	hypothetical prote
28	296.5	18.7	299	2 T22706	hypothetical prote
29	296.5	18.7	1690	1 CGHUIB	collagen alpha 4(I

30	294	18.6	1843	2 S18803	collagen alpha 1(V
31	293.5	18.6	1049	1 CGBOIS	collagen alpha 1(I
32	290.5	18.4	633	2 B40983	collagen alpha 1(X
33	289.5	18.3	281	2 T32765	hypothetical prote
34	289.5	18.3	1492	2 A40333	collagen alpha 1(I
35	288.5	18.2	1024	2 S18251	collagen alpha 1(X
36	288.5	18.2	1486	1 B40333	collagen alpha 1(I
37	288	18.2	1838	1 CGHUIB	collagen alpha 1(V
38	287.5	18.2	300	2 T24482	hypothetical prote
39	287.5	18.2	319	2 T32250	hypothetical prote
40	287.5	18.2	920	2 B34493	collagen alpha 1(I
41	287.5	18.2	2944	2 S18251	collagen alpha 1(V
42	286.5	18.1	1466	1 CGHUIB	collagen alpha 1(I
43	286	18.1	1487	1 CGHUIB	collagen alpha 1(I
44	285.5	18.0	140	2 A05249	collagen alpha 1(I
45	285.5	18.0	299	2 T29956	hypothetical prote
46	285.5	18.0	1464	2 S59856	collagen alpha 1(I
47	285	18.0	298	2 UC1448	collagen col-34 -
48	285	18.0	886	2 I50694	collagen alpha 1(I
49	285	18.0	1418	2 T45467	collagen alpha 1(I
50	285	18.0	1691	1 S22937	collagen alpha 5(I
51	283.5	17.9	310	2 S10696	collagen alpha 1(I
52	283	17.9	1419	2 A41182	collagen alpha 1(I
53	283	17.9	1487	2 B41182	collagen alpha 1(I
54	282	17.8	295	2 T22833	hypothetical prote
55	282	17.8	673	1 CGBOIS	collagen alpha 1(I
56	282	17.8	888	2 S28791	collagen alpha 1(X
57	282	17.8	1142	2 UX0369	collagen alpha 1(X
58	282	17.8	1806	1 CGHUIB	collagen alpha 1(X
59	281.5	17.8	688	2 A53330	collagen alpha 2(I
60	281	17.8	453	2 S18804	collagen alpha 2(I
61	281	17.8	964	1 CGCHIS	collagen alpha 2(I
62	281	17.8	1388	2 A53317	collagen alpha 1(X
63	280.5	17.7	283	2 T19141	hypothetical prote
64	280.5	17.7	615	2 A05269	collagen alpha 1(I
65	280.5	17.7	636	2 S41067	collagen alpha 1(I
66	280	17.7	1747	2 A45974	collagen alpha 1(X
67	279.5	17.7	921	2 S40495	collagen alpha 1(I
68	278.5	17.6	201	2 T46404	hypothetical prote
69	278.5	17.6	1028	1 CGHUIB	collagen alpha 1(V
70	278	17.6	1366	2 T18763	collagen alpha 2(I
71	277.5	17.5	324	2 CGHUIB	hypothetical prote
72	277.5	17.5	488	2 A27353	collagen alpha 1(I
73	277.5	17.5	920	2 A45748	collagen alpha 1(I
74	277.5	17.5	1707	2 A35526	collagen alpha 2(I
75	277	17.5	296	2 T21070	hypothetical prote
76	277	17.5	1603	2 S23810	collagen alpha 1(X
77	276.5	17.5	286	2 S34665	collagen, cuticula
78	276.5	17.5	299	2 T25407	hypothetical prote
79	276	17.4	1019	1 A32856	collagen alpha 1(V
80	276	17.4	1373	1 A43281	collagen alpha 2(I
81	276	17.4	1414	1 S23809	collagen alpha 2(I
82	275.5	17.4	302	2 A31921	collagen dpv-13 pr
83	275	17.4	1497	2 T49607	procollagen type V
84	275	17.4	1758	2 T29350	hypothetical prote
85	274.5	17.4	623	2 A45137	collagen alpha 4(I
86	274.5	17.4	730	2 A36226	collagen alpha 1 c
87	274.5	17.4	921	2 S42617	collagen alpha 1(I
88	274.5	17.4	1669	1 CGMS48	collagen alpha 1(I
89	274	17.3	324	2 T33149	hypothetical prote
90	273.5	17.3	324	2 A31920	collagen sgc-1 pre
91	273.5	17.3	438	2 S53787	collagen alpha cha
92	273.5	17.3	1546	1 CGHUIB	collagen alpha 2(X
93	273	17.3	283	2 US0170	collagen col-19 -
94	273	17.3	289	2 T27708	hypothetical prote
95	273	17.3	290	2 A32249	collagen - sea urc
96	273	17.3	341	2 T16296	hypothetical prote
97	273	17.3	366	2 S11449	collagen short cha
98	273	17.3	918	2 S23377	collagen alpha 2(V
99	273	17.3	1022	2 S04111	collagen alpha 2(V
100	273	17.3	1670	1 CGHUIB	collagen alpha 3(I

## ALIGNMENTS

## RESULT 1

CGR010  
collagen alpha 1(X) chain precursor - human  
N/Alternate names: procollagen alpha 1(X) chain  
C/Species: Homo sapiens (man)  
C/Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text change 09-Jul-2004  
C/Accession: S26396; S30086; S15826; S18249; A43901; I51870; S21856  
R/Reichemberger, B.; Beler, F.; Luvallie, P.; Olsen, B.R.; von der Mark, K.; Bertling, W.  
FEBS Lett. 311, 305-310, 1992  
A/Title: Genomic organization and full-length cDNA sequence of human collagen X.  
A/Reference number: S26396; MUID:93012005; PMID:1397333  
A/Accession: S26396  
A/Molecule type: DNA  
A/Residues: 1-680 <RE1>  
A/Cross-references: UNIPROT:Q03692; EMBL:X68952; EMBL:X72578; EMBL:X72579; EMBL:X72580; R/Apte, S.S.  
submitted to the EMBL Data Library, March 1992  
A/Reference number: S30085  
A/Accession: S30086  
A/Molecule type: DNA  
A/Residues: 'TTPYGVWCWCLL', 52-680 <APT>  
A/Cross-references: EMBL:X65120; NID:923129  
A/Note: the initial difference is probably due to translation of an intronic sequence  
R/Apte, S.; Mettel, M.G.; Olsen, B.R.  
FEBS Lett. 282, 393-396, 1991  
A/Title: Cloning of human alpha-1(X) collagen DNA and localization of the COL10A1 gene  
A/Reference number: S15826; MUID:91243838; PMID:2037056  
A/Accession: S15826  
A/Molecule type: DNA  
A/Residues: 561-647 'G', 649-666 <AP2>  
A/Cross-references: EMBL:X56879; NID:930013; PIDN:CAA41686.1; PID:930014  
R/Thomas, J.T.; Cresswell, C.J.; Raehn, B.; Nicolai, H.; Jones, T.; Solomon, E.; Grant, M.  
Biochem. J. 280, 617-623, 1991  
A/Title: The human collagen X gene. Complete primary translated sequence and chromosomal  
A/Reference number: S18249; MUID:92109659; PMID:1764025  
A/Accession: S18249  
A/Molecule type: DNA  
A/Residues: 1-26, 'T', 28-680 <THO>  
A/Cross-references: EMBL:X60382; NID:930094; PIDN:CAA42933.1; PID:930095  
A/Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 532-Ala  
R/Reichemberger, E.; Aigner, T.; von der Mark, K.; Stoss, H.; Bertling, W.  
Dev. Biol. 148, 562-572, 1991  
A/Title: In situ hybridization studies on the expression of type X collagen in fetal human  
A/Reference number: A43901; MUID:92077285; PMID:1743401  
A/Accession: A43901  
A/Molecule type: mRNA  
A/Residues: 547-656 <RE2>  
A/Cross-references: GB:M74050; GB:D57494; NID:9339884; PIDN:AAA61221.1; PID:9553796  
A/Note: sequence extracted from NCBI backbone (NCBI:65012, NCBI:65014)  
R/Mallis, G.A.; Raehn, B.; Sweetman, W.A.; Thomas, J.T.; Super, M.; Evans, G.; Grant, M.E.  
Am. J. Hum. Genet. 54, 169-178, 1994  
A/Title: Amino acid substitutions of conserved residues in the carboxyl-terminal domain  
pe Schmid.  
A/Reference number: I51870; MUID:94136476; PMID:8304336  
A/Accession: I51870  
A/Status: translated from GB/EMBL/DBDJB  
A/Molecule type: mRNA  
A/Residues: 520-597, 'D', 599-680 <NAL>  
A/Cross-references: GB:S68511; NID:9545180; PIDN:AAC60615.1; PID:9545181  
A/Note: mutant sequence from patient with metaphyseal chondrodysplasia type Schmid  
A/Note: a second mutant sequence with 614-Pro is also described  
C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
ed and subsequently O-glycosylated.  
C/Genetics:  
A/Gene: GDB:COL10A1  
A/Cross-references: GDB:128635; OMIM:120110  
A/Map position: 6q21-6q22  
A/Intons: 52/1  
A/Note: a defect in this gene may cause Schmid metaphyseal chondrodysplasia  
C/Complex: type X collagen may be a homotrimer

## C/Function:

A/Description: structural component of extracellular fibrous polymer specifically and tre  
be important for skeletogenesis  
C/Superfamily: collagen alpha 1(VIII) chain, complement C1q carboxyl-terminal homology  
C/Keywords: coiled coll; extracellular matrix; glycoprotein; homotrimer; hydroxyllysine, t  
F/1-18/Domain: signal sequence #status predicted <SIG>  
F/19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>  
F/19-56/Domain: amino-terminal nonhelical #status predicted <NC2>  
F/520-519/Region: interrupted helical  
F/520-680/Domain: amino-terminal nonhelical #status predicted <NC1>  
F/553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>  
F/617/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.6%; Score 516.5; DB 1; Length 680;  
Best Local Similarity 41.5%; Pred. No. 5,8e-30;  
Matches 120; Conservative 29; Mismatches 93; Indels 47; Gaps 8;

Qy 26 DFRKSPQVLCSLDPGQPP--GPGAPSGSGMGMGPFPGKDGQDGDHGRDGSSEEG 82  
Db 397 DGRKGNP---GLPGKGDPGVGPPGLPGVPAGAKGPHGNHGEAGPRGAPGIPGTRG 452  
Qy 83 PPRGTGNRKGKPKGKGAIGRAGPRG--PKGVNTPGKKGTPGKKGKKGEPGLRGP 140  
Db 453 PIGPPGIPGPPGSKGDPSPGPPGAIATKGLNGPT--GPPGPPRGHSGBPGLGPP 509  
Qy 141 -----CSCSGHT---KSAFSAVATKSPYRRLPT 167  
Db 510 PGPPPGQAVMPEGRIKAGRPISLSTPLVSNQVITGMVAFYILSKAPALGTPI 569  
Qy 168 KPDRIINMGSGHYNASSGKFCVCGPIYFTYDITLANKRLAIGLVHNGQYRIRTPDANT 227  
Db 570 PFDKILVNRQOHDPRTGIFTCQIPGIYYSYVHVHVGTHVWGLYKNGTPWVYTDYDT 629  
Qy 228 -GNHDAVSSGTTIALKQGDVWLCIYFSEONGHFDYPTWDSLFTGFLI 275  
Db 630 KGYLDASGSALTIDLTENDQWMLPVAESNGIYSSEY-VHSSFSGFLV 677

## RESULT 2

S31216

collagen alpha 1(X) chain precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text change 09-Jul-2004

C/Accession: S31216; S28807; S22215; S30127; I48299; S26397; S31830

R/Kong, R.Y.C.; Kwan, K.M.; Lau, E.T.; Thomas, J.T.; Boot-Handford, R.P.; Grant, M.E.; Cf

Eur. J. Biochem. 213, 99-111, 1993

A/Title: Intron-exon structure, alternative use of promoter and expression of the mouse c

A/Reference number: S31216; MUID:93238750; PMID:8477738

A/Accession: S31216

A/Molecule type: DNA

A/Residues: 1-680 &lt;KON&gt;

A/Cross-references: UNIPROT:Q05306; EMBL:Z21610; NID:949793; PIDN:CAA47736.1; PID:949794

R/Bilma, K.; Berola, I.; Rosati, R.; Metcraeranta, M.; Garofalo, S.; Peraelae, M.; de Cron

Biochem. J. 289, 247-253, 1993

A/Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and exp

A/Reference number: S28807; MUID:93143676; PMID:8424763

A/Accession: S28807

A/Molecule type: DNA

A/Residues: 1-285, 'A', 287-680 &lt;ELI&gt;

A/Cross-references: EMBL:X67348; NID:950480; PIDN:CAA47763.1; PID:950481

R/Bilma, K.; Metcraeranta, M.; Kallio, J.; Peraelae, M.; Berola, I.; Garofalo, S.; de Cron

Biochim. Biophys. Acta 1130, 78-80, 1992

A/Title: Specific hybridization probes for mouse alpha-2(I)X and alpha-1(X) collagen mRN

A/Reference number: S22215; MUID:92182017; PMID:1543751

A/Accession: S22215

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 385-450, 'K', 452-627 &lt;ELA&gt;

A/Cross-references: EMBL:X63013; NID:949795; PIDN:CAA44741.1; PID:949796

R/Apte, S.S.; Olsen, B.R.

Matrix 13, 165-179, 1993

A/Title: Characterization of the mouse type X collagen gene.

A/Reference number: S30127; MUID:93261348; PMID:8492743

A:Accession: S30127  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-12, 'F', 14-26, 'S', 28-247, 'L', 249-285, 'A', 287-305, 'P', 307-416, 'S', 418-499, 'L', 501-566, 'C', 568, 'V', 570-600, 'L', 602-603, 'S', 605-606, 'L', 608-609, 'S', 611-612, 'L', 614-615, 'S', 617-618, 'L', 620-621, 'S', 623-624, 'L', 626-627, 'S', 629-630, 'L', 632-633, 'S', 635-636, 'L', 638-639, 'S', 641-642, 'L', 644-645, 'S', 647-648, 'L', 650-651, 'S', 653-654, 'L', 656-657, 'S', 659-660, 'L', 662-663, 'S', 665-666, 'L', 668-669, 'S', 671-672, 'L', 674-675, 'S', 677-678, 'L', 680-681, 'S', 683-684, 'L', 686-687, 'S', 689-690, 'L', 692-693, 'S', 695-696, 'L', 698-699, 'S', 701-702, 'L', 704-705, 'S', 707-708, 'L', 710-711, 'S', 713-714, 'L', 716-717, 'S', 719-720, 'L', 722-723, 'S', 725-726, 'L', 728-729, 'S', 731-732, 'L', 734-735, 'S', 737-738, 'L', 740-741, 'S', 743-744, 'L', 746-747, 'S', 749-750, 'L', 752-753, 'S', 755-756, 'L', 758-759, 'S', 761-762, 'L', 764-765, 'S', 767-768, 'L', 770-771, 'S', 773-774, 'L', 776-777, 'S', 779-780, 'L', 782-783, 'S', 785-786, 'L', 788-789, 'S', 791-792, 'L', 794-795, 'S', 797-798, 'L', 800-801, 'S', 803-804, 'L', 806-807, 'S', 809-810, 'L', 812-813, 'S', 815-816, 'L', 818-819, 'S', 821-822, 'L', 824-825, 'S', 827-828, 'L', 830-831, 'S', 833-834, 'L', 836-837, 'S', 839-840, 'L', 842-843, 'S', 845-846, 'L', 848-849, 'S', 851-852, 'L', 854-855, 'S', 857-858, 'L', 860-861, 'S', 863-864, 'L', 866-867, 'S', 869-870, 'L', 872-873, 'S', 875-876, 'L', 878-879, 'S', 881-882, 'L', 884-885, 'S', 887-888, 'L', 890-891, 'S', 893-894, 'L', 896-897, 'S', 899-900, 'L', 902-903, 'S', 905-906, 'L', 908-909, 'S', 911-912, 'L', 914-915, 'S', 917-918, 'L', 920-921, 'S', 923-924, 'L', 926-927, 'S', 929-930, 'L', 932-933, 'S', 935-936, 'L', 938-939, 'S', 941-942, 'L', 944-945, 'S', 947-948, 'L', 950-951, 'S', 953-954, 'L', 956-957, 'S', 959-960, 'L', 962-963, 'S', 965-966, 'L', 968-969, 'S', 971-972, 'L', 974-975, 'S', 977-978, 'L', 980-981, 'S', 983-984, 'L', 986-987, 'S', 989-990, 'L', 992-993, 'S', 995-996, 'L', 998-999, 'S', 1000-1001, 'L', 1002-1003, 'S', 1005-1006, 'L', 1008-1009, 'S', 1011-1012, 'L', 1014-1015, 'S', 1017-1018, 'L', 1020-1021, 'S', 1023-1024, 'L', 1026-1027, 'S', 1029-1030, 'L', 1032-1033, 'S', 1035-1036, 'L', 1038-1039, 'S', 1041-1042, 'L', 1044-1045, 'S', 1047-1048, 'L', 1050-1051, 'S', 1053-1054, 'L', 1056-1057, 'S', 1059-1060, 'L', 1062-1063, 'S', 1065-1066, 'L', 1068-1069, 'S', 1071-1072, 'L', 1074-1075, 'S', 1077-1078, 'L', 1080-1081, 'S', 1083-1084, 'L', 1086-1087, 'S', 1089-1090, 'L', 1092-1093, 'S', 1095-1096, 'L', 1098-1099, 'S', 1100-1101, 'L', 1102-1103, 'S', 1105-1106, 'L', 1108-1109, 'S', 1111-1112, 'L', 1114-1115, 'S', 1117-1118, 'L', 1120-1121, 'S', 1123-1124, 'L', 1126-1127, 'S', 1129-1130, 'L', 1132-1133, 'S', 1135-1136, 'L', 1138-1139, 'S', 1141-1142, 'L', 1144-1145, 'S', 1147-1148, 'L', 1150-1151, 'S', 1153-1154, 'L', 1156-1157, 'S', 1159-1160, 'L', 1162-1163, 'S', 1165-1166, 'L', 1168-1169, 'S', 1171-1172, 'L', 1174-1175, 'S', 1177-1178, 'L', 1180-1181, 'S', 1183-1184, 'L', 1186-1187, 'S', 1189-1190, 'L', 1192-1193, 'S', 1195-1196, 'L', 1198-1199, 'S', 1200-1201, 'L', 1202-1203, 'S', 1205-1206, 'L', 1208-1209, 'S', 1211-1212, 'L', 1214-1215, 'S', 1217-1218, 'L', 1220-1221, 'S', 1223-1224, 'L', 1226-1227, 'S', 1229-1230, 'L', 1232-1233, 'S', 1235-1236, 'L', 1238-1239, 'S', 1241-1242, 'L', 1244-1245, 'S', 1247-1248, 'L', 1250-1251, 'S', 1253-1254, 'L', 1256-1257, 'S', 1259-1260, 'L', 1262-1263, 'S', 1265-1266, 'L', 1268-1269, 'S', 1271-1272, 'L', 1274-1275, 'S', 1277-1278, 'L', 1280-1281, 'S', 1283-1284, 'L', 1286-1287, 'S', 1289-1290, 'L', 1292-1293, 'S', 1295-1296, 'L', 1298-1299, 'S', 1300-1301, 'L', 1302-1303, 'S', 1305-1306, 'L', 1308-1309, 'S', 1311-1312, 'L', 1314-1315, 'S', 1317-1318, 'L', 1320-1321, 'S', 1323-1324, 'L', 1326-1327, 'S', 1329-1330, 'L', 1332-1333, 'S', 1335-1336, 'L', 1338-1339, 'S', 1341-1342, 'L', 1344-1345, 'S', 1347-1348, 'L', 1350-1351, 'S', 1353-1354, 'L', 1356-1357, 'S', 1359-1360, 'L', 1362-1363, 'S', 1365-1366, 'L', 1368-1369, 'S', 1371-1372, 'L', 1374-1375, 'S', 1377-1378, 'L', 1380-1381, 'S', 1383-1384, 'L', 1386-1387, 'S', 1389-1390, 'L', 1392-1393, 'S', 1395-1396, 'L', 1398-1399, 'S', 1400-1401, 'L', 1402-1403, 'S', 1405-1406, 'L', 1408-1409, 'S', 1411-1412, 'L', 1414-1415, 'S', 1417-1418, 'L', 1420-1421, 'S', 1423-1424, 'L', 1426-1427, 'S', 1429-1430, 'L', 1432-1433, 'S', 1435-1436, 'L', 1438-1439, 'S', 1441-1442, 'L', 1444-1445, 'S', 1447-1448, 'L', 1450-1451, 'S', 1453-1454, 'L', 1456-1457, 'S', 1459-1460, 'L', 1462-1463, 'S', 1465-1466, 'L', 1468-1469, 'S', 1471-1472, 'L', 1474-1475, 'S', 1477-1478, 'L', 1480-1481, 'S', 1483-1484, 'L', 1486-1487, 'S', 1489-1490, 'L', 1492-1493, 'S', 14

C;Genetics:

A:Gene: COL10A1

C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

C:Keywords: coiled coil, extracellular matrix; glycoprotein; homotrimer

F:1-18/Domain: signal sequence #status predicted <Sig>

F:19-674/Product: collagen alpha 1(X) chain #status predicted <Mat>

F:547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 31.8%; Score 502.5; DB 2; Length 674;  
Best Local Similarity 41.8%; Pred. No. 5,9e-29;  
Matches 117; Conservative 28; Mismatches 94; Indels 41; Gaps 8;

Oy 29 KSPDVLCSLPGGP--GPPCAPSGSMGRMGPGKDGDDGDGSDSGEPPG 85  
|||::|  
Db 400 KGNP----GLPFGXDPGIAGSPGLPCPVPAKAAGVPHNGNCAGRGVIGTGTPIG 455  
Oy 86 RTGNRGKRGPKRGAAGARGAPRRPGK--VNTRPGKHGTPTKKGPBGKGEPLRP--- 140  
|||::|  
Db 456 PPGIPPGPSKGD--VTGPQPAPGIAVKANGLTPGPPGPPGPNABEPGLPGPBP 512  
|||::|  
Oy 141 -----SCSGSHT--KSAFVAATKSYPERRLRPIRKDKILMNE 176  
|||::|  
Db 513 PGPPGYALBEDPYKAGQRPFVSANGCVTMPPASFTVLSTKAVPALGPTIPDDKLILYNK 572  
|||::|  
Oy 177 GGHNASSGKFVCVGVPETIYFTDTLANGLAIGLVHNCQYRIKTFPDAN-TGNHDVASG 235  
|||::|  
Db 573 QOHDPRTGTFTCKIPETIYFSYHIHWKGTHAWGVLYKNGTPWMTYYDEXIKGYLDQASG 632  
|||::|  
Oy 236 STIALKQGDEVMQLIFYSEONGLFYDPVYTDSLFTGPILI 275  
|||::|  
Db 633 SAVDLTENDQVWLQLPNAGSNGL-YSPFYVHSSFSGLV 671  
|||::|

RESULT 4  
S23297  
collagen alpha 1(X) chain precursor - chicken  
N:Alternate names: type X collagen  
C:Species: Gallus gallus (chicken)  
C:Date: 07-Oct-1994 #sequence, revision 10-Nov-1995 #text, change 09-Jul-2004  
A:Accession: S23297; S55594; S77711; I50218  
R:Ninomiya, Y.; Castagnola, F.; Gerecke, D.; Gordon, M.K.; Jacekko, O.; LuValle, P.; MCC  
Magnuchi, N.; Olsen, B.R.  
A:Title: The molecular biology of collagens with short triple-helical domains.  
A:Reference number: S22243  
A:Accession: S23297  
A:Status: preliminary  
A:Molecule type: MRNA  
A:Residues: 1-674 <NIN>  
A:CROSS-references: UNIPROT:P08125  
R:LuValle, P.; Ninomiya, Y.; Rosenblum, N.D.; Olsen, B.R.  
J. Biol. Chem. 263, 18378-18385, 1988  
A:Title: The type X collagen gene. Intron sequences split the 5'-untranslated region and  
A:Reference number: A31896; MUID:89054019; PMID:2461368  
A:Accession: A31896  
A:Molecule type: MRNA  
A:Residues: 1-75 <LUV>  
R:Ninomiya, Y.; Gordon, M.; van der Rest, M.; Schmid, T.; Linsenmayer, T.; Olsen, B.R.  
J. Biol. Chem. 261, 5041-5050, 1986  
A:Title: The developmentally regulated type X collagen gene contains a long open reading  
A:Reference number: I50218; MUID:8618227; PMID:3082876  
A:Accession: S65594  
A:Molecule type: DNA  
A:Residues: 104-112, 'X', 114-117;453-466 <NIN2>  
C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology  
C:Keywords: coiled coil, extracellular matrix; glycoprotein; homotrimer; hydroxyproline;  
F:1-18/Domain: signal sequence #status predicted <Sig>  
F:547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>



omains similar to those of type X collagen.

A:Reference number: A34246; MUID:89380199; PMID:2476437

A:Accession: A34246

A:Molecule type: mRNA

A:Residues: 1-744 <YAM>

A:Cross-references: UNIPROT:P14282; GB:J05042; NID:G164895; PID:AAA1204.1; PID:G164896

C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>

F:21-744/Region: amino-terminal nonhelical

F:118-571/Region: interrupted helical

F:572-744/Region: carboxyl-terminal nonhelical

F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 27.7%; Score 439; DB 1; Length 744;

Best Local Similarity 31.8%; Pred. No. 2.6e-24;

Matches 113; Conservative 33; Mismatches 99; Indels 110; Gaps 9;

30 GSPOLVCSLPGPGPP-----GPGAPGSGMGMGMPGPKD---GDDGH 71

390 GEPGLP-GIPGMPGPAIGPPGKSGIVGPGPPGKBPGLQGPGRGLGEGVP 448

72 DGRDGSSEGGPPGRTGNRGP-----GPKGK-----AGA 101

449 PGIRGLPGPIGPKGBAGHKLPGIPGVPLGPGKEPGIPDDGLQGPPIGRTGPGP 508

102 IGRAGPPGPKVNGTPGK-----HGTGPKGPKGKKGEPGLP----- 140

509 IGPPIGPKGEPGLPBPFPVGVGKPGVAGLHGPGRKGLGPGQGPGLPBPFP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGAPVEMPAFTALPTAPP 628

163 ERLPIKDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRIT 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGAGVWVALFKNNBPVWT 688

223 FDA-NTGNHVAASSTIALKQGBEWMQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLDQAGSAAVLLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

RESULT 8

S23298

collagen alpha 1(VIII) chain - chicken

C:Species: Gallus gallus (chicken)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: S23298

R:Kinomura, Y.; Casdagli, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; Mc

maguchi, N.; Olsen, B.R.

In Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pre

A:Title: The molecular biology of collagens with short triple-helical domains.

A:Reference number: S22243

A:Accession: S23298

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-744 <NIN>

A:Cross-references: UNIPROT:Q7LZR2

C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 27.5%; Score 435; DB 1; Length 744;

Best Local Similarity 31.8%; Pred. No. 5e-24;

Matches 113; Conservative 32; Mismatches 100; Indels 110; Gaps 9;

30 GSPOLVCSLPGPGPP-----GPGAPGSGMGMGMPGPKD---GDDGH 71

390 GEPGLP-GIPGMPGPAIGPPGKSGIVGPGPPGKBPGLQGPGRGLGEGVP 448

72 DGRDGSSEGGPPGRTGNRGP-----GPKGK-----AGA 101

449 PGIRGLPGPIGPKGBAGHKLPGIPGVPLGPGKEPGIPDDGLQGPPIGRTGPGP 508

102 IGRAGPPGPKVNGTPGK-----HGTGPKGPKGKKGEPGLP----- 140

509 IGPPIGPKGEPGLPBPFPVGVGKPGVAGLHGPGRKGLGPGQGPGLPBPFP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGAPVEMPAFTALPTAPP 628

163 ERLPIKDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRIT 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGAGVWVALFKNNBPVWT 688

223 FDA-NTGNHVAASSTIALKQGBEWMQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLDQAGSAAVLLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

RESULT 9

JC4708

gelatin-binding 28k protein precursor - human

N:Alternate names: adipose specific collagen-like factor

C:Species: Homo sapiens (man)

C>Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 09-Jul-2004

C:Accession: JC4708; JC4944

R:Maeda, K.; Okubo, K.; Shimomura, I.; Funahashi, T.; Matsuzawa, Y.; Matsubara, K.

Biochem. Biophys. Res. Commun. 221, 286-289, 1996

A:Title: cDNA cloning and expression of a novel adipose specific collagen-like factor, a

A:Reference number: JC4708; MUID:96224171; PMID:8619847

A:Accession: JC4708

A:Molecule type: mRNA

A:Residues: 1-244 <MAE>

A:Cross-references: UNIPROT:Q15848; DDBJ:D45371; NID:9871886; PID:BA08227.1; PID:98718

A:Experimental source: adipose tissue

R:Nakano, Y.; Tohe, T.; Choi-Miura, N.H.; Maeda, T.; Tomita, M.

J. Biochem. 120, 803-812, 1996

A:Title: Isolation and characterization of GBP28, a novel gelatin-binding protein purified

A:Reference number: JC4944; MUID:97103474; PMID:8947845

A:Accession: JC4944

A:Molecule type: protein

A:Residues: 19-38;93-100;101-112;135-149;173-178 <NAK>

A:Comment: This protein is an endogenous factor that binds with a collagen-like domain.

C:Gene: apm1

C:Keywords: adipose tissue; glycoprotein; hydroxyproline

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-244/Product: gelatin-binding 28kDa protein #status experimental <MAT>

F:112-107/Region: collagen-like

F:95/Modified site: 4-hydroxyproline (Pro) #status experimental

F:230/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 27.3%; Score 431.5; DB 2; Length 244;

Best Local Similarity 40.2%; Pred. No. 2.8e-24;

Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

30 GSPOLVCSLPG-----PGP-----GPGAPGSGMGMGMPGPKDGDGDDSDG 79

5 GAVLLALPLPHDDPTTQGGVLLPLPKGA--CTGMVA--GIPGHGHN----- 50

80 BEGPPGRTGNRKGKGPGRKGAIGRAGPRGKGVNPGKNGTGPGRKGGKGPGLPG 139

51 --GAPRGDGRDGTGKGBKDDPLGPKGIGTGVPGABGPPGPGIQRKSPGE-- 106

140 PCSGSGHTKSAFSAVATKSYRERLPIKDKILMNEGHNASSGKFCVGPPIYFTY 199

107 ----GAVVYSAFVGLFTYTIINMIRFTKIFYNQNMNDSTGKFNCHIRGLYFYAY 162

200 DITLANKHLAGLVHNGQYRITFDA-NTGNHVAASSTIALKQGBEWMQITF-YSEON 257

163 HITVMDQVKSLEPKOKAMLFYDQYQENNVDAAGSVALHLEVDQVWLQVYGEGERN 222

258 GLFYDPYWTDSLFTGFLIYAD 278





## RESULT 13

S29328 Complement subcomponent C1q chain C - mouse

C/Species: Mus musculus (house mouse)

C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C/Accession: S29328

R/Petry, F.; Reid, K.B.M.; Loos, M.

Biochem. J. 209, 129-134, 1992

A/Title: Isolation, sequence analysis and characterization of cDNA clones coding for the ecebelin.

A/Reference number: S29328; MUID:9301118; PMID:1396691

A/Accession: S29328

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-246 &lt;PRT&gt;

A/Cross-references: UNIPROT:Q02105; EMBL:X66295; NID:G50228; PIDN:CAA4693.1; PID:G50225  
C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom F/122-245/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 24.8%; Score 393; DB 2; Length 246;

Best Local Similarity 38.8%; Pred. No. 1.7e-21;

Matches 94; Conservative 36; Mismatches 84; Indels 28; Gaps 11;

Qy 42 QGPPGPGAPGAPSGMGMGPGKQDGDHGDSDGSEGGPRTGNGKPGKXKA 101

Db 28 QASAGCGTIGMPCGM-----PGAPGKDGHGLOGKPGEPITPAVPGTGKQKQKGP 81

Qy 102 IGRAGPAPGKGVNTPGKGTGPKGKGGKGGPGLPGPCSCSGHTKSAFVA-VTKSY 160

Db 82 PGNHKGKNGPRTS-----GLPGDPGPRGPGEGVEGRYK--QKH-QSVPTVTRQTQY 132

Qy 161 PREHLPIKPKILMBEGHYNASSGKVCVPGIYFTYDT-LAN--KHLAIGLVNNGQ 217

Db 133 PEANALRFNSVVTNPGSHVPSGTGTCVPLGYFTYSHYANLCHLNLMLA---- 188

Qy 218 YRIKTF-DANTGNMDVAGSTTLAKOGDEWVLOIFYSEONGLFYDPTWDTLFTGFLIY 276

Db 189 -RVASFCDHMFNSKQVSSGGLALRLQKQDEWVLSV--NDYNGM-VGIEGSSVSFGFLIF 244

Qy 277 AD 278

Db 245 PD 246

## RESULT 14

C1HUQB Complement subcomponent C1q chain B precursor [validated] - human

N/Alternate names: complement subcomponent C1q beta chain

C/Species: Homo sapiens (man)

C/Date: 22-May-1991 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004

C/Accession: B23422; B90304; A90301; B90315; A03206

R/Reid, K.B.M.

Biochem. J. 231, 729-735, 1985

A/Title: Molecular cloning and characterization of the complementary DNA and gene coding A/Reference number: A23422; MUID:86076906; PMID:3000358

A/Accession: B23422

A/Molecule type: DNA

A/Residues: 'HS', 1-32 &lt;REI&gt;

A/Cross-references: UNIPROT:P02746

A/Note: the authors translated the codon AGT for the second position as Arg; they were u

A/Accession: A23422

A/Molecule type: mRNA

A/Residues: 28-253 &lt;REI&gt;

A/Cross-references: EMBL:X03084

A/Note: the authors translated the codon ACA for residue 46 as Ile

R/Reid, K.B.M.

Biochem. J. 179, 367-371, 1979

A/Title: Complete amino acid sequences of the three collagen-like regions present in sub

A/Reference number: A90304; MUID:80020137; PMID:486087

A/Accession: B90304

A/Molecule type: protein

A/Residues: 'E', 29-84, 'D', 86-99, 'P', 101-135 &lt;RES&gt;

R/Reid, K.B.M.; Thompson, E.O.P.

Biochem. J. 173, 863-868, 1978

A/Title: Amino acid sequence of the N-terminal 108 amino acid residues of the B chain of

A/Reference number: A90301; MUID:79041552; PMID:708376

A/Accession: A90301

A/Molecule type: protein

A/Residues: 28-99, 'P', 101-195 &lt;RES&gt;

R/Reid, K.B.M.; Gagnon, J.; Frampton, J.

Biochem. J. 203, 559-569, 1982

A/Title: Completion of the amino acid sequences of the A and B chains of subcomponent C1

A/Reference number: A90315; MUID:8283890; PMID:6961411

A/Accession: B90315

A/Molecule type: protein

A/Residues: 136-253 &lt;RE4&gt;

A/Note: 176-Glx may also be present

C/Comment: The first component of complement is a calcium-dependent complex of the three  
activation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.  
C/Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfide  
(see PIR:C1HUQ) chain. Equimolar amounts of the A, B, and C chains are found after redu

C/Genetics:

A/Genes: GDB:C1QB

A/Cross-references: GDB:119043; OMIM:120570

A/Map position: 1p36.3-1p34.1

C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom

C/Keywords: complement pathway; glycoprotein; heterodimer; hydroxylysine; hydroxyproline

F/1-27/Domain: signal sequence #status predicted &lt;SIG&gt;

F/28-253/Product: complement subcomponent C1q chain B #status experimental &lt;MAT&gt;

F/33-116/Domain: collagenase, triple helix &lt;COL&gt;

F/123-249/Domain: complement C1q carboxyl-terminal homology &lt;C1Q&gt;

F/28/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status experimen

F/31/Disulfide bonds: interchain (to chain A-26) #status experimental

F/35, 38, 41, 53, 56, 65, 83, 86, 101, 104, 107/Modified site: 4-hydroxyproline (Pro) #status expe

F/59, 62, 77, 92, 98, 110/Modified site: 5-hydroxylysine (Lys) #status experimental

F/59, 62, 98, 110/Binding site: carboxylate (Lys) (covalent) #status experimental

Query Match 23.8%; Score 376.5; DB 1; Length 253;

Best Local Similarity 37.2%; Pred. No. 2.8e-20;

Matches 94; Conservative 34; Mismatches 92; Indels 33; Gaps 10;

Qy 33 QLVCS-LPGPGPPGAPGAPSGMGMGPGKQDGDHGDSDGSEGGPRTGNGK 91

Db 28 QLSCTGPPAIPGIGTPTGPDGPPTGPKGKGGPGLGADHGEKEDPSPGPG 87

Qy 92 KPGPKKAGATIGRAGPAPGKGVNTPGKGTGPKGKGGKGGPGLPGPCSCSGHTKSA 151

Db 88 KVGPK-----GPGPK-----GAPAPGAPGKSGD-----YATQKIA 123

Qy 152 PSVAVTKSY-REHLPIKPKILMBEGHYNASSGKVCVPGIYFTYDTTLANKHLAI 210

Db 124 PSATRTINVLPRBDQTRFDHVTNNMNNYEPKFTCKVPGIYFTYHNS-SRGNLCV 182

Qy 211 GLVNGQ-YRIKTF-DANTGNMDVAGSTTLAKOGDEWVLOIFYSEONGLFYDPTWDT 267

Db 183 NLMGRRAQKVCVPCVAYNTFQVTTGGVWLKLEGBENFLQ--TDKSLT-GEAGAN 239

Qy 268 SLTGFLLVADQD 280

Db 240 SIFSGFLFPDME 252

## RESULT 15

S19018 Complement subcomponent C1q chain A precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004

C/Accession: S19018

R/Petry, F.; Reid, K.B.M.; Loos, M.

J. Immunol. 147, 3988-3993, 1991

A/Title: Gene expression of the A- and B-chain of mouse C1q in different tissues and the

A/Reference number: S19018; MUID:92043789; PMID:1940381

A/Accession: S19018

A/Molecule type: mRNA

A/Residues: 1-245 &lt;PRT&gt;

A/Cross-references: UNIPROT:P98086; EMBL:X58861; NID:G50226; PIDN:CAA41664.1; PID:G50227

C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom  
F,116-243/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 23.3%; Score 369; DB 2; Length 245;

Best Local Similarity 37.1%; Pred. No. 9.4e-20; Matches 92; Conservative 33; Mismatches 87; Indels 36; Gaps 9;

```
QY 35 VCSLP-GPGPPGPPGAPSPGMMGRMGFPKDG-QDGHDDRGDSGEGPPGRTGNRGK 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 VCRAPNCKDGAPEGNPGRRPGLKGERGEPGAAGIRTGIRGPKDPPESGPPGKRNVL 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 93 PGPKGAKAIGRAGPRGPKGVNGTPGKHGTPGKKGPKKGEPGLPGPCSCSGHTYSAP 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 PGP---SGPLGDSGPGGLKGVKGNPG-----NIRDQ-----RPAF 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 153 SVAVTKSYPRERLPRIKEDKILMNEGHNASSGKFCVCGVPGIYVETVD-ITLANKHLAIG 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 S-AIRQNPMTLGNVVIIDKVLTNQESPPYQNHTRGFICAVPGFYFNFQVLSKMDLCLFIK 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 LVHNGQYRIRTFDANTGN--HDVASGSTILALKQDEVMLOIFYSQNGLFYDPYWTDS 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 SSSGGQPRDLSFENTNNKGLFQVLAGTVIQLRRGDVWIE--KDPKGRRIYGTEDADS 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 269 LFTGFLIY 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 IFSGFLIF 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: December 25, 2004, 08:39:35  
Job time : 76.0538 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: December 25, 2004, 06:42:17 / Search time 422.991 Seconds  
(without alignments)  
387.672 Million cell updates/sec

Title: US-10-621-787-2

Perfect score: 1582

Sequence: 1 MIPWVLACALPCADPLG.....TDSLFTGFLIYADDDNEV 285

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1582	100.0	285	1	COT2_HUMAN
2	1492	94.3	294	2	Q8DBT4
3	898.5	56.8	312	2	Q8CHX9
4	895	56.6	289	1	CQT7_HUMAN
5	889	56.2	289	2	Q8BVD7
6	521.5	33.0	333	2	Q8IUD4
7	521.5	33.0	333	2	AAQ88512
8	516.5	32.6	680	1	CA1A_HUMAN
9	511	32.3	680	1	CA1A_MOUSE
10	503	31.8	295	2	Q9ZIK4
11	502.5	31.8	674	1	CA1A_BOVIN
12	502.5	31.8	675	2	Q9N178
13	493	31.2	674	1	CA1A_CHICK
14	482	30.5	508	1	COT1_ONCOK
15	457.5	28.9	419	1	COLE_LEPMA
16	444	28.1	703	1	CA28_HUMAN
17	442.5	28.0	743	1	CA18_MOUSE
18	442.5	28.0	744	2	Q8BGL6
19	442.5	28.0	744	2	Q92188
20	441.5	27.9	244	2	Q6QW57
21	441.5	27.9	244	2	AA667924
22	440	27.8	640	2	Q6KAQ4
23	440	27.8	640	2	BAD21403
24	440	27.8	699	2	Q6PIC4
25	440	27.8	699	2	AA65148
26	439	27.7	744	1	CA18_RABIT
27	436	27.6	744	1	CA18_HUMAN
28	436	27.6	744	2	AA688919
29	435	27.5	744	2	Q71Z82
30	431.5	27.3	244	1	APM1_HUMAN
31	426	26.9	247	2	Q8BRW2

32	425	26.9	240	2	Q95MO4	Q95M4 bos taurus
33	425	26.9	247	1	APM1_MOUSE	Q6Q994 mus musculus
34	424	26.8	195	2	Q8BZ53	Q8BZ53 mus musculus
35	424	26.8	247	2	AAH28770	AAH28770 mus musculus
36	420	26.5	243	2	Q95JD7	Q95JD7 macaca mulia
37	419.5	26.5	235	2	Q6DJ45	Q6DJ45 xenopus tro
38	413	26.1	244	2	Q8K3R4	Q8K3R4 rattus norv
39	411.5	26.0	235	2	Q6GLS9	Q6GLS9 xenopus lae
40	409	25.9	243	2	Q6PP07	Q6PP07 sus scrofa
41	409	25.9	243	2	RAT00459	RAT00459 sus scrofa
42	408	25.8	242	2	Q6Q2K6	Q6Q2K6 sus scrofa
43	408	25.8	242	2	AA575592	AA575592 sus scrofa
44	408	25.8	243	2	Q8K479	Q8K479 mus musculus
45	404	25.5	243	2	Q8R002	Q8R002 mus musculus
46	403	25.5	243	1	COT5_HUMAN	Q9BX10 homo sapien
47	402	25.4	243	2	Q8N6P2	Q8N6P2 homo sapien
48	400	25.3	246	2	Q6DI63	Q6DI63 mus musculus
49	393	24.8	246	1	C1CQ_MOUSE	Q02105 mus musculus
50	392	24.8	243	2	Q7YRF8	Q7YRF8 sus scrofa
51	392	24.8	244	2	Q76C76	Q76C76 canis famli
52	392	24.8	244	2	BAD15362	BAD15362 canis fam
53	376.5	23.8	251	1	C1QB_HUMAN	P02746 homo sapien
54	369	23.3	245	1	C1QA_MOUSE	P98086 mus musculus
55	369	23.3	245	2	Q9DCM6	Q9DCM6 m mus muscu
56	369	23.3	253	2	AAH67001	AAH67001 mus muscu
57	367.5	23.2	230	2	Q6LA35	Q6LA35 lampetra ja
58	367	23.2	236	2	Q6GNJ3	Q6GNJ3 xenopus lae
59	366.5	23.2	194	2	Q95J95	Q95J95 canis famli
60	366	23.1	245	1	C1QC_HUMAN	P02747 homo sapien
61	363.5	23.0	258	1	C1RF_MOUSE	Q08992 mus musculus
62	362	22.9	253	1	C1QB_MOUSE	P14106 mus musculus
63	357	22.6	245	2	Q7Z502	Q7Z502 homo sapien
64	357	22.6	253	1	C1QB_RAT	P31721 rattus norv
65	355.5	22.5	258	1	C1RF_HUMAN	Q75973 homo sapien
66	355	22.4	247	2	Q6GPX5	Q6GPX5 xenopus lae
67	348.5	22.0	245	1	C1QA_HUMAN	P02745 homo sapien
68	341	21.6	255	1	GLIC_MOUSE	Q9684 mus musculus
69	331	20.9	287	2	Q8CFR0	Q8CFR0 mus musculus
70	328	20.7	238	2	Q86Z23	Q86Z23 homo sapien
71	327.5	20.7	246	2	Q6NZU2	Q6NZU2 brachydanio
72	327.5	20.7	246	2	AAH65968	AAH65968 brachydanio
73	316	20.0	287	2	Q7Z5L3	Q7Z5L3 homo sapien
74	315.5	19.9	250	2	Q7Z282	Q7Z282 brachydanio
75	311.5	19.7	1069	2	Q6LA88	Q6LA88 homo sapien
76	311.5	19.7	1069	2	CA67261	CA67261 homo sapi
77	311.5	19.7	1461	1	CA11_CANPA	Q9X8J7 canis famli
78	311.5	19.7	1461	2	Q76045	Q76045 homo sapien
79	311.5	19.7	1464	1	CA11_HUMAN	P02452 homo sapien
80	311.5	19.7	1464	2	Q8N473	Q8N473 homo sapien
81	310	19.6	1225	2	Q6PCU3	Q6PCU3 mus musculus
82	310	19.6	1225	2	AAH59281	AAH59281 mus muscu
83	310	19.6	1453	1	CA11_MOUSE	P11087 mus musculus
84	310	19.6	1453	2	Q810J9	Q810J9 mus musculus
85	309.5	19.6	187	2	Q6L7J6	Q6L7J6 cyprinus ca
86	309.5	19.6	187	2	BAD22535	BAD22535 cyprinus
87	309.5	19.6	671	1	CA11_RAT	P02454 rattus norv
88	308	19.5	1453	2	Q63079	Q63079 rattus norv
89	306	19.3	215	1	HP25_TAMSI	Q06576 tamias sibi
90	306	19.3	1453	1	CA11_CHICK	P02457 gallus galli
91	305.5	19.3	246	1	Q9ES30	Q9ES30 m collageneo
92	304	19.2	246	1	COT3_HUMAN	Q9BX14 homo sapien
93	304	19.2	246	2	BAC1676	BAC1676 homo sapi
94	303.5	19.2	304	2	AAK6452	AAK6452 caenorhabd
95	303.5	19.2	431	2	Q95Y22	Q95Y22 caenorhabd
96	303.5	19.2	779	1	CA11_BOVIN	P02455 bos taurus
97	302.5	19.1	343	2	Q46388	Q46388 equus cabal
98	300.5	19.0	297	2	Q8MXR1	Q8MXR1 caenorhabd
99	300.5	19.0	299	2	Q22393	Q22393 caenorhabd
100	300.5	19.0	717	2	Q9NQ52	Q9NQ52 homo sapien

## ALIGNMENTS

RESULT 1  
COT2\_HUMAN STANDARD; PRT; 265 AA.  
ID COT2\_HUMAN STANDARD; PRT; 265 AA.  
AC Q9BJU5;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Complement C1q tumor necrosis factor-related protein 2 precursor  
DE (UNQ6349/PRO21054).  
GN Name=C1QTNF2; Synonyms=CTRP2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Piddington C.S., Bishop P.;  
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.",  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=22867296; PubMed=12975309; DOI=10.1101/gr.1293003;  
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hase P.E., Heldens S.,  
RA Hang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
RA Lewis L., Liao D., Mark M., Robble E., Sanchez C., Schoenfeld J.,  
RA Shendgiri S., Simmons L., Singh J., Smith V., Stinson U., Vagstad A.,  
RA Vanden R., Watnabe C., Wleand D., Woods K., Xie M.-H., Yansura D.,  
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Zoddard A., Wood W.I.,  
RA Godowski P., Gray A.;  
RT "The secreted protein discovery initiative (SPDI), a large-scale  
RT effort to identify novel human secreted and transmembrane proteins: a  
RT bioinformatics assessment.";  
RL Genome Res. 13:2265-2270(2003).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA TISSUE=Brain, and Muscle;  
RC MEDLINE=22368257; PubMed=12477937; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Whiting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA TISSUE=PROT entry is copyright. It is produced through a collaboration  
RA between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
RA the European Bioinformatics Institute. There are no restrictions on its  
RA use by non-profit institutions as long as its content is in no way  
RA modified and this statement is not removed. Usage by and for commercial  
RA entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
RA or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AF329836; AAKI7960.1; -  
CC EMBL; AY358839; AAO89198.1; -  
CC

DR EMBL; BC011699; AAH11699.1; -  
DR EMBL; BC054506; AAH54506.1; -  
DR HGSP; Q60994; 1C28.  
DR Genew; HGNC:14325; C1QTNF2.  
DR InterPro; IPR001073; C1q.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR008983; TNF\_1like.  
DR Pfam; PF00386; C1q; 1.  
DR Pfam; PF01391; Collagen; 2.  
DR PRINTS; PR000007; COMPLEMENTC1Q.  
DR SMART; SM00110; C1Q; 1.  
DR PROSITE; PS01113; C1q; 1.  
KW Collagen; Signal.  
FT SIGNAL 1 15 Potential.  
FT CHAIN 16 285 Complement C1q tumor necrosis factor-  
FT DOMAIN 40 141 related protein 2.  
FT DOMAIN 143 285 C1q.  
SQ SEQUENCE 285 AA; 29952 MW; 7E31FF986BD4EDFA CRC64;  
Query Match 100.0%; Score 1582; DB 1; Length 285;  
Best Local Similarity 100.0%; Pred. No. 8.9e-98;  
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MIPVVLALCALPCADPDLGAFARDFRKSPOLVCSLPDPGPPGAPGSGMMGRM 60  
DB 1 MIPVVLALCALPCADPDLGAFARDFRKSPOLVCSLPDPGPPGAPGSGMMGRM 60  
QY 61 GFPGKQGDHDDHDDRGSGEGSPGRGNGNGKPGPKAKAIGRAGRGRGVNGTGK 120  
DB 61 GFPGKQGDHDDHDDRGSGEGSPGRGNGNGKPGPKAKAIGRAGRGRGVNGTGK 120  
QY 121 GFPGKKGPKKPKGAPGLPGPCSGSGHTKSAFSAVATKSPREBLPKFKILMNEG 180  
DB 121 GFPGKKGPKKPKGAPGLPGPCSGSGHTKSAFSAVATKSPREBLPKFKILMNEG 180  
QY 181 MASSGKFCVCGVPPIYFTYDITLANKLAIGLVHNGGYRIRTPDANTGNHDVAGST 240  
DB 181 MASSGKFCVCGVPPIYFTYDITLANKLAIGLVHNGGYRIRTPDANTGNHDVAGST 240  
QY 241 LKQGDVWLQIFYSQNGLFYDPYWDLSFTGFLIYADQDDPNEV 285  
DB 241 LKQGDVWLQIFYSQNGLFYDPYWDLSFTGFLIYADQDDPNEV 285  
RESULT 2  
Q9DBU4 PRELIMINARY; PRT; 294 AA.  
ID Q9DBU4 PRELIMINARY; PRT; 294 AA.  
AC Q9DBU4;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
DE Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched  
DE library, clone:181003K05 product:COMPLEMENT-C1Q TUMOR NECROSIS  
DE FACTOR-RELATED PROTEIN (COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-RELATED  
DE PROTEIN 2) homolog (C1qtnf2 protein).  
GN Name=C1qtnf2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=C57BL/6J; TISSUE=Pancreas;  
RC MEDLINE=99279253; PubMed=10349636;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Pancreas;  
RX MEDLINE=21085660; PubMed=11217851;  
RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).

RL [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RA The PANTOM Consortium,

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs." Nature 420:563-573(2002).

RL [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RX MEDLINE=20499374; PubMed=11042159;

RA Kinnari P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kanno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes." Genome Res. 10:1617-1630(2000).

RL [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Kanno H., Akiyama J., Nishi K., Kitamura T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaethiwa K., Fujikake S., Inoue K., Togawa Y., Izawa M., Obara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "Riken integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer." Genome Res. 10:1757-1771(2000).

RL [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RA Adachi J., Aizawa K., Akahira S., Akiyama T., Arai A., Aono H., Atakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;

RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RL [7]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N-3;

RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., Brownstein M.J., Uedin T.B., Tishiyuki S., Carninci P., Prange C., Rana S.S., Loughlan N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McWeeney P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Abbey J., Helton B., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Kozminski M.I., Skelton U., Small D.E., Scherch A., Schein J.E., Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP [8]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N-3;

RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.

RA Strausberg R.;

RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; AK007683; BAB25187.1; -.

DR EMBL; BC030324; AAH03024.1; -.

DR HSSP; Q60994; 1C28.

DR MGD; MGI:1916433; C1qtnf2.

DR GO; GO:0005737; C:cytoplasm; IEA.

DR GO; GO:0006817; P:phosphate transport; IEA.

DR InterPro; IPR01073; C1q.

DR InterPro; IPR008160; C1qgen.

DR InterPro; IPR008983; TWF-like.

DR Pfam; PF00386; C1q; 1.

DR Pfam; PF01391; Collagen; 2.

DR PRINTS; PR00007; COMPLEMENTC1Q.

DR SMART; SM00110; C1Q; 1.

DR PROSITE; PS01113; C1Q; 1.

KW COLLAGEN.

SQ SEQUENCE 294 AA; 30865 MW; 6D3905AE7C19E6FA CRC64;

Query Match 94.3%; Score 1492; DB 2; Length 294;

Best Local Similarity 94.4%; Pred. No. 9.1e-92;

Matches 269; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Oy 1 MIPVTLALCALPCADPLLGAFAARDPRKSPQLVCSLPFGQGPFGAPGPGMGGRM 60

Db 10 MISVTLALCALPCADPLMLGAFARDFKGGPQLVCSLPFGQGPFGAPGPGMGGRM 69

Oy 61 GFGPKDQDGDHGDGSDGSEBPPGRTGNRKGPKPKAGAIAGAPRPGKVGNGTPGKH 120

Db 70 GFGPKDQDGDGDGRGDSGSEBPPGRTGNRKGPKPKAGAIAGAPRPGVSGTPGKH 129

Oy 121 GTPPKKPKPKKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 180

Db 130 GTPPKKPKPKKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 189

Oy 181 NASGKGFVCGVPGVGYFTYDITLANKLAIGLVNGQVRIPTPDANTGNHVASGSTLLA 240

Db 190 NASGKGFVCGVPGVGYFTYDITLANKLAIGLVNGQVRIPTPDANTGNHVASGSTLLA 249

Oy 241 LKQDDEVWLTQIFVSEQNGLFYDPYMTDSLFTGFLIYADQDDPNEV 285

Db 250 LKQDDEVWLTQIFVSEQNGLFYDPYMTDSLFTGFLIYADQDDPNEV 294

RESULT 3

O8CHX9 PRELIMINARY; PRT; 312 AA.

AC O8CHX9;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE C1qntf7 protein (Fragment).

GN Name=C1qntf7;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Mammary gland;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,

RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schermer A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC038308; AAH38308.1; -  
 DR HSSP: Q60994; 1C28.  
 DR MGD; MGI:1925911; Clqtnf7.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro: IPR001073; Clq.  
 DR InterPro: IPR008160; Collagen.  
 DR InterPro: IPR008983; TNF-like.  
 DR Pfam; PF00386; Clq; 1.  
 DR Pfam; PF01391; Collagen; 2.  
 DR PRINTS; PR00007; COMPLEMENTC1Q.  
 DR SMART; SM00110; Clq; 1.  
 DR PROSITE; PS01113; Clq; 1.  
 KM Collagen.  
 FT NON TMR.  
 SQ SEQUENCE 312 AA; 33162 MW; 1E240A83A5014F43 CRC64;  
 Query Match 56.6%; Score 898.5; DB 2; Length 312;  
 Best Local Similarity 62.7%; Pred. No. 3.4e-52;  
 Matches 163; Conservative 32; Mismatches 62; Indels 3; Gaps 1;  
 QY 24 RDRPRK---SPQVCSLPDPGPPGPPGAGPSGMMGRMGFPKQDQDHDGRGSGE 80  
 DB 42 RANQAKESYSPRYICSPGLPPGPPGANGSPGPHRGILPRDRDRGKGGKGGKGT 101  
 QY 81 EGPPGRTNGKPKPKKAGAGIAGRPGRGVNGTGKGTGPKKGPCKKGGKGGPGL 140  
 DB 102 AGLKGTGPPGLAGKEDQGTGKGGPIGPGEGKEGVGPACPPPKKDRGQDPPGLPGV 161  
 QY 141 CSCSGHTKSAFSAVATKSYPRERLPFKDKILNBSGHNVAASGKTFVCGVPGIYFTYD 200  
 DB 162 CRGSIVLKSAFSGITTSYERELPIIFNKVLFNBSGHNVPATGKFLCAPPGIYFTSYD 221  
 QY 201 ITLANKLAIGLVNCGYRITFPDANTGNHDVAGSTITLAKQDEWVLOIFYEONGLP 260  
 DB 222 ITLANKLAIGLVNCGYRITFPDANTGNHDVAGSTIVLYLPDEWVLEIFNDONGLP 281  
 QY 261 YDPYWTDSLFTGFLIYADQD 280  
 DB 282 SDPGWADSLFSGFLIYDVT 301

OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Piddington C.S., Sheppard P.O., Bishop P., Laesser G.W.,  
 RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.",  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.F., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buétow K.H., Schaefer C.M., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusita K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- SIMILARITY: Contains 1 C1q domain.  
 CC -1- SIMILARITY: Contains 1 collagenous domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; AF329839; AAKI7963.1; -  
 DR EMBL; BC022187; AAH22187.1; -  
 DR EMBL; BC024015; AAH24015.1; -  
 DR HSSP; Q60994; 1C28.  
 DR Genew; HGNC:14362; C1QTNF7.  
 DR InterPro: IPR001073; Clq.  
 DR InterPro: IPR008160; Collagen.  
 DR InterPro: IPR008983; TNF-like.  
 DR Pfam; PF00386; Clq; 1.  
 DR Pfam; PF01391; Collagen; 2.  
 DR PRINTS; PR00007; COMPLEMENTC1Q.  
 DR SMART; SM00110; Clq; 1.  
 DR PROSITE; PS01113; Clq; 1.  
 KM Collagen; Signal.  
 FT SIGNAL 1 16 Potential.  
 FT CHAIN 17 289 Complement C1q tumor necrosis factor-  
 FT DOMAIN 38 139 related protein 7.  
 FT DOMAIN 141 276 Collagen-like.  
 SQ SEQUENCE 289 AA; 30683 MW; A61609FF86D26946 CRC64;  
 Query Match 56.6%; Score 895; DB 1; Length 289;  
 Best Local Similarity 63.6%; Pred. No. 5.4e-52;  
 Matches 159; Conservative 33; Mismatches 58; Indels 0; Gaps 0;  
 QY 31 SPQVCSLPDPGPPGPPGAGPSGMMGRMGFPKQDQDHDGRGSGEGPPGRTGNR 90  
 DB 29 SPRTICSPGLPPGPPGANGSPGPHRGILPRDRDRGKGGKGGKGTAGLRGKTGPL 88  
 QY 91 GKPPKKAAGIAGRAGRGKGVNGTTPGKGTGKPKKGGKGGKGGPGLPGKSCSGSHTKS 150  
 DB 89 GLAGEKDQGTGKGGPIGPGEGKEGVGPPIGPPKDRGBQDPPGLPGVCRCSIVLKS 148

QY 151 AFSVAVTKSPRERLPKEDKILNMGHVNASSGKPGVCGIYYFTYDITLANKHLAI 210  
 DB 149 AFSVGITTSYBERPLIFENKVLFNEGHYNPATGKPLCAPPGIYFSDITLANKHLAI 208  
 QY 211 GLVHNGQYRITFPANTGNHVDVAGSSTIIALKQGDVWLIQIFYSBONGLFYDPYWTDSL 270  
 DB 209 GLVHNGQYRITFPANTGNHVDVAGSSTIYIQLPDEVWLEIFFTDQNLFSDPGADSLF 268  
 QY 271 TGFLLIYADQD 280  
 DB 269 SGFLLYVDT 278  
 RESULT 5  
 Q8BVD7 PRELIMINARY; PRT; 289 AA.  
 ID Q8BVD7  
 AC Q8BVD7;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Mus musculus 16 days embryo lung cDNA, RIKEN full-length enriched library, clone:8430425G24 product:COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-RELATED PROTEIN homolog.  
 GN Name=C1qtnf7;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Mech. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA The FANTOM Consortium;  
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team; Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Komano H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Azawa K., Nagaoka S., Sasaki N., Carninci P., Komano H., Akiyama J., Nishi K., Kitanai T., Tashiro H., Itoh M., Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsuura H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA Adachi J., Azawa K., Akahira S., Akimura T., Aono H., Arai A., Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hasegawa T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Komano H., Kouda M., Koya S., Kurihara C., Matsuura T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;  
 RA Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.  
 RL EMBL; AK078818; BAC37409.1; -.  
 DR HSSP; Q60994; IC28.  
 DR MGD; MGI:1925911; C1qtnf7.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR001073; C1q.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008983; TNF\_like.  
 DR Pfam; PF00386; C1q\_1.  
 DR Pfam; PF01391; Collagen\_2.  
 DR PRINTS; PRO0007; COMPLEMENTC1Q.  
 DR SMART; SM00110; C1Q\_1.  
 DR KEGG; collagen.  
 SQ SEQUENCE 289 AA; 30483 MW; 514AC72C318832A CRC64;  
 Query Match 56.2%; Score 889; DB 2; Length 289;  
 Best Local Similarity 63.6%; Pred. No. 1.3e-51;  
 Matches 159; Conservative 31; Mismatches 60; Indels 0; Gaps 0;  
 QY 31 SPOLVCSLPQPGPGPGAPGAPSGMGMGMFPGKDGQDGHGSGEGPPTGNR 90  
 DB 29 SPRYICSTPLGPRPPGPGANGSPGPHRIGLPGRDGDGRKSGEGGTGLKGTPL 88  
 QY 91 GKPPKGAAGIAGPRPGKGVNGTPGKGTGPKGPKGKGBGLPFGSCSGGHTKS 150  
 DB 89 GLAGEKDGQSTGKGPFGPEGKEVGVPAPPGPKGRDQDGLPGVCRCSIVLKS 148  
 QY 151 AFSVAVTKSPRERLPKEDKILNMGHVNASSGKPGVCGIYYFTYDITLANKHLAI 210  
 DB 149 AFSVGITTSYBERPLIFENKVLFNEGHYNPATGKPLCAPPGIYFSDITLANKHLAI 208  
 QY 211 GLVHNGQYRITFPANTGNHVDVAGSSTIIALKQGDVWLIQIFYSBONGLFYDPYWTDSL 270  
 DB 209 GLVHNGQYRITFPANTGNHVDVAGSSTIYIQLPDEVWLEIFFTDQNLFSDPGADSLF 268  
 QY 271 TGFLLIYADQD 280  
 DB 269 SGFLLYVDT 278  
 RESULT 6  
 Q8IUU4 PRELIMINARY; PRT; 333 AA.  
 ID Q8IUU4  
 AC Q8IUU4;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-OCT-2004 (TREMBLrel. 28, Last sequence update)  
 DE Hypothetical protein WGC48915 (R1W6503).  
 GN Name=WGC48915; ORFNames=UNQ6503;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932;

RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marzetta K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scalapeto M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrincci P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywnicki M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences." J.  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RN SEQUENCE FROM N.A.  
 KC TISSUE=Pancreas;  
 RL Straubeberg R.;  
 RN Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22887296; PubMed=12975309;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Batton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Helens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seehagiri S., Simmons L., Singh V., Smith V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RT "The secreted protein discovery initiative (SPDI), a large-scale  
 RT effort to identify novel human secreted and transmembrane proteins: a  
 RT bioinformatics assessment." J.  
 RL Genome Res. 13:2265-2270(2003).  
 RX EMBL; BC040438; AA040438.1; -;  
 DR EMBL; AY358145; AA088512.1; -;  
 DR HSSP; Q60994; 1C28.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR001073; Clq.  
 DR InterPro; IPR008161; Clq\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008983; TWF\_like.  
 DR Pfam; PF00386; Clq; 1.  
 DR Pfam; PF01391; Collagen; 3.  
 DR PRINTS; PR00007; COMPLEMENT1Q.  
 DR ProDom; PD000007; Clq\_helix; 1.  
 DR PROSITE; PS01113; Clq; 1.  
 KW Collagen; Hypothetical protein.  
 SQ SEQUENCE 333 AA; 34649 MW; E9EBF803A034E307 CRC64;

Query Match 33.0%; Score 521.5; DB 2; Length 333;  
 Best Local Similarity 38.3%; Pred. No. 4,7e-27;  
 Matches 128; Conservative 37; Mismatches 102; Indels 67; Gaps 11;

QY 4 WLLALCALPCADBLGAFARRD-FRKGSPTLVCS-----LPGQGPFGPAGPAGPSGM 57  
 DB 5 WLLALTEI-CT-----GNINSQDTCRQGHFGIPGNPGHNLPGRDGDAKGDAGDEP 58

QY 58 GRMGFPKGDQDGDHGDGRGDS-----GEE 81  
 DB 59 GRPSPKDGSGSKGRGADGKVEAKIGKDGQSRGSPKGNPKGLAGPWEKGLGET 118

QY 82 GPGRTGNRKPGRPKGKAGAIGRAGPRGPKGVNGTPGKNGTPGK--GPKKGGEGGLP 138  
 DB 119 GPGQKGNKNDVGPTEGPRGNIGPLGPTGLPQPMGPIGKPGKGAEGPTGPGSEGV 178

QY 139 GPGSCGS-----GHT-----KSAPSAVT--KSYPRERLPFKDKILNMEGHYNASS 184  
 DB 179 GIRKWKDGRGKRGKIGTTLVLPKSAFTVGLTVLSKSPSDVPIKFDKILNERNHYDTAA 238

QY 185 GKFCVGVPIYFTYDITLANKHLAIGLVNNGOYRIRTPDANTGNHDVAGSTTLAKOG 244  
 DB 239 GKFTCHLAGYVYFTYHTVPSRVNQLVKNQVILHTKDAVMSSEDOASGIVLQKLG 298

QY 245 DEWVLQIFYSBQ-NGLFYDPYRWDSLFTGFLIYA 277  
 DB 299 DEWVLQVTGGERFNGFLPADBD-DTPTTGLFLFS 331

RESULT 7  
 ID AA088512 PRELIMINARY; PRT; 333 AA.  
 AC AA088512;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE R1W6503.  
 GN UN06503.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primata; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12975309;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Batton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Helens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seehagiri S., Simmons L., Singh V., Smith V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RT "The Secreted Protein Discovery Initiative (SPDI), a Large-Scale  
 RT Effort to Identify Novel Human Secreted and Transmembrane Proteins: A  
 RT Bioinformatics Assessment." J.  
 RL Genome Res. 13:2265-2270(2003).  
 DR EMBL; AY358145; AA088512.1; -;  
 SQ SEQUENCE 333 AA; 34649 MW; E9EBF803A034E307 CRC64;

Query Match 33.0%; Score 521.5; DB 2; Length 333;  
 Best Local Similarity 38.3%; Pred. No. 4,7e-27;  
 Matches 128; Conservative 37; Mismatches 102; Indels 67; Gaps 11;

QY 4 WLLALCALPCADBLGAFARRD-FRKGSPTLVCS-----LPGQGPFGPAGPAGPSGM 57  
 DB 5 WLLALTEI-CT-----GNINSQDTCRQGHFGIPGNPGHNLPGRDGDAKGDAGDEP 58

QY 58 GRMGFPKGDQDGDHGDGRGDS-----GEE 81  
 DB 59 GRPSPKDGSGSKGRGADGKVEAKIGKDGQSRGSPKGNPKGLAGPWEKGLGET 118

QY 82 GPGRTGNRKPGRPKGKAGAIGRAGPRGPKGVNGTPGKNGTPGK--GPKKGGEGGLP 138  
 DB 119 GPGQKGNKNDVGPTEGPRGNIGPLGPTGLPQPMGPIGKPGKGAEGPTGPGSEGV 178

QY 139 GPGSCGS-----GHT-----KSAPSAVT--KSYPRERLPFKDKILNMEGHYNASS 184  
 DB 179 GIRKWKDGRGKRGKIGTTLVLPKSAFTVGLTVLSKSPSDVPIKFDKILNERNHYDTAA 238

QY 185 GKFCVGVPIYFTYDITLANKHLAIGLVNNGOYRIRTPDANTGNHDVAGSTTLAKOG 244  
 DB 239 GKFTCHLAGYVYFTYHTVPSRVNQLVKNQVILHTKDAVMSSEDOASGIVLQKLG 298

QY 245 DEWVLQIFYSBQ-NGLFYDPYRWDSLFTGFLIYA 277  
 DB 299 DEWVLQVTGGERFNGFLPADBD-DTPTTGLFLFS 331



## RESULT 8

CA1A\_HUMAN STANDARD; PRT; 680 AA.  
 ID CA1A\_HUMAN  
 AC 003692;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Collagen alpha 1(X) chain precursor.  
 GN Name=COL10A1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92109659; PubMed=1764025;  
 RA Thomas J.T., Cresswell C.J., Raab B., Nicolai H., Jones T.,  
 RA Solomon E., Grant M.E., Boot-Handford R.P.;  
 RT "The human collagen X gene. Complete primary translated sequence and  
 RT chromosomal localization.";  
 RL Biochem. J. 280:617-623(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93012005; PubMed=1397333;  
 RA Reichenberger E., Beier F., Luvallie P., Olsen B.R., von der Mark K.,  
 RA Bertling W.M.;  
 RT "Genomic organization and full-length cDNA sequence of human collagen  
 RT X.";  
 RL FEBS Lett. 311:305-310(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Beier F., Lamm M.B., von der Mark K.;  
 RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Williams S.;  
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 52-680 FROM N.A.  
 RX MEDLINE=92267014; PubMed=1587271;  
 RA Ade S.S., Seldin M.F., Hayaishi M., Olsen B.R.;  
 RT "Cloning of the human and mouse type X collagen genes and mapping of  
 RT the mouse type X collagen gene to chromosome 10.";  
 RL Eur. J. Biochem. 206:217-224(1992).  
 RN [6]  
 RP SEQUENCE OF 561-666 FROM N.A.  
 RX MEDLINE=91243838; PubMed=2037056;  
 RA Ade S., Matrei M.-G., Olsen B.R.;  
 RT "Cloning of human alpha 1(X) collagen DNA and localization of the  
 RT COL10A1 gene to the q21-q22 region of human chromosome 6.";  
 RL FEBS Lett. 282:393-396(1991).  
 RN [7]  
 RP SEQUENCE OF 547-655 FROM N.A.  
 RX MEDLINE=92077285; PubMed=1743401;  
 RA Reichenberger E., Aigner T., von der Mark K., Stoeb H., Bertling W.;  
 RT "In situ hybridization studies on the expression of type X collagen in  
 RT fetal human cartilage.";  
 RL Dev. Biol. 148:562-572(1991).  
 RN [8]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=97255959; PubMed=9101290;  
 RA Kivivand H., Tromp G., Prockop D.J.;  
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-  
 RT associated collagen (type IX), and network-forming collagen (type X)  
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";  
 RL Hum. Mutat. 9:300-315(1997).  
 RN [9]  
 RP VARIANTS SMCD ASP-598 AND PRO-614.  
 RX MEDLINE=94134476; PubMed=8304336;  
 RA Wallis G.A., Raab B., Sweetman W.A., Thomas J.T., Super M., Evans G.,  
 RA Grant M.E., Boot-Handford R.P.;

RT "Amino acid substitutions of conserved residues in the carboxyl-  
 RT terminal domain of the alpha 1(X) chain of type X collagen occur in  
 RT two unrelated families with metaphyseal chondrodysplasia type  
 RT Schmid.";  
 RL Am. J. Hum. Genet. 54:169-178(1994).  
 RN [10]  
 RP VARIANTS SMCD ARG-591.  
 RX MEDLINE=94272470; PubMed=8004099;  
 RA McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francomano C.A.;  
 RT "Additional mutations of type X collagen confirm COL10A1 as the Schmid  
 RT metaphyseal chondrodysplasia locus.";  
 RL Hum. Mol. Genet. 3:303-307(1994).  
 RN [11]  
 RP VARIANTS SMCD VAL-618.  
 RX MEDLINE=95181449; PubMed=7876225;  
 RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;  
 RT "Type X collagen multimer assembly in vitro is prevented by a Gly618  
 RT to Val mutation in the alpha 1(X) NCI domain resulting in Schmid  
 RT metaphyseal chondrodysplasia.";  
 RL J. Biol. Chem. 270:4558-4562(1995).  
 RN [12]  
 RP VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.  
 RX MEDLINE=95331767; PubMed=7607655;  
 RA Bonaventure J., Chamande F., Maroteaux P.;  
 RT "Mutations in three subdomains of the carboxy-terminal region of  
 RT collagen type X account for most of the Schmid metaphyseal  
 RT dysplasias.";  
 RL Hum. Genet. 96:58-64(1995).  
 RN [13]  
 RP VARIANTS SMCD PRO-600.  
 RX MEDLINE=96375754; PubMed=8782043;  
 RA Wallis G.A., Raab B., Sykes B., Bonaventure J., Maroteaux P.,  
 RA Zabel B., Wyne-Davies R., Grant M.E., Boot-Handford R.P.;  
 RT "Mutations within the gene encoding the alpha 1(X) chain of type X  
 RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but  
 RT not several other forms of metaphyseal chondrodysplasia.";  
 RL J. Med. Genet. 33:450-457(1996).  
 RN [14]  
 RP VARIANTS SMCD GLU-18 AND ARG-18.  
 RX MEDLINE=97220591; PubMed=9067753;  
 RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;  
 RT "Mutations in the N-terminal globular domain of the type X collagen  
 RT gene (COL10A1) in patients with Schmid metaphyseal chondrodysplasia.";  
 RL Hum. Mutat. 9:131-135(1997).  
 RN [15]  
 RP VARIANTS SMD GLU-595.  
 RX MEDLINE=99057503; PubMed=9837818;  
 RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,  
 RA Nakamura Y.;  
 RT "Mutation of the type X collagen gene 'COL10A1' causes  
 RT spondylometaphyseal dysplasia.";  
 RL Am. J. Hum. Genet. 63:1659-1662(1998).  
 RN [16]  
 RP VARIANTS SMCD CYS-597.  
 RX MEDLINE=99069781; PubMed=9852679;  
 RA Sawai H., Ida A., Nakata Y., Koyama K.;  
 RT "Novel missense mutation resulting in the substitution of tyrosine by  
 RT cysteine at codon 597 of the type X collagen gene associated with  
 RT Schmid metaphyseal chondrodysplasia.";  
 RL J. Hum. Genet. 43:259-261(1998).  
 CC -1- FUNCTION: Type X collagen is a product of hypertrophic  
 CC chondrocytes and has been localized to presumptive  
 CC mineralization zones of hyaline cartilage.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- PMT: Prolines at the third position of the tripeptide repeating  
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
 CC -1- DISBASR: Defects in COL10A1 are the cause of Schmid type  
 CC metaphyseal chondrodysplasia (SMCD) [MIM:156500]. SMCD is a  
 CC dominantly inherited disorder of the osseous skeleton. The  
 CC cardinal features of the phenotype are mild short stature, coxa  
 CC vara and a waddling gait. Radiography usually shows sclerosis of  
 CC the ribs, flaring of the metaphyses, and a wide irregular growth  
 CC plate, especially of the knees.

```

CC -1- DISEASE: Defects in COL10A1 are a cause of spondylometaphyseal
CC dysplasia Japanese type (SMD) [MIM:120110]. SMD comprises a
CC heterogeneous group of heritable skeletal dysplasias characterized
CC by modifications of the vertebral bodies of the spine and
CC metaphyses of the tubular bones.
CC -1- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60382; CAA42933.1; -
DR EMBL; X65120; CAA46236.1; -
DR EMBL; X98568; CAA67178.1; -
DR EMBL; AL121963; CAB87590.1; -
DR EMBL; S68531; AAC60615.1; -
DR EMBL; M74050; AAA61686.1; -
DR EMBL; X72579; CAA51170.1; -
DR EMBL; X72580; CAA51170.1; JOINED.
DR PIR; S26396; CGHUID.
DR PDB; 1GR3; X-ray; A=521-680.
DR Genew; HGNC:2185; COL10A1.
DR MIM; 120110; -
DR MIM; 156500; -
DR MIM; 184250; -
DR GO; GO:0005581; C:collagen; TAS.
DR GO; GO:0001501; P:skeletal development; TAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR009883; TNF-like.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 8.
DR PRINTS; PRO0007; COMPLEMENTC1Q.
DR PROSITE; PS01113; C1Q; 1.
KW 3D-structure; Collagen; Connective tissue; Disease mutation;
KW Extracellular matrix; Hydroxylation; Polymorphism; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 680 Collagen alpha 1(X) chain.
FT DOMAIN 19 56 Nonhelical region (NC2).
FT DOMAIN 57 519 Triple-helical region.
FT DOMAIN 520 680 Nonhelical region (NC1).
FT DOMAIN 545 680 C1q.
FT VARIANT 18 18 G->E (in SMD).
FT VARIANT 18 18 /FTID=VAR_001838.
FT VARIANT 18 18 G->R (in SMD).
FT VARIANT 18 18 /FTID=VAR_001839.
FT VARIANT 18 18 G->R (in dbSNP:2228547).
FT VARIANT 545 545 /FTID=VAR_001840.
FT VARIANT 591 591 C->R (in SMD).
FT VARIANT 591 591 /FTID=VAR_001841.
FT VARIANT 595 595 G->E (in SMD and SMD).
FT VARIANT 597 597 /FTID=VAR_001842.
FT VARIANT 597 597 Y->C (in SMD).
Query Match 32.6%; Score 516.5; DB 1; Length 680;
Best Local Similarity 41.5%; Pred. No. 2,1e-26;
Matches 120; Conservative 29; Mismatches 93; Indels 47; Gaps 8;

```

```

Db 510 PGPPGPGQAVMEGFIKAGQRPSTPLVSANQVGTGVPASFTVILSKAPALGTPPI 569
Qy 168 KPDIIIMNEGHYNASGKRGVGVPGIYYFTYITLANKHLAGLVHNGQYRIRTPDANT 227
Db 570 PFDILYNRQOHDPRTGITTQIPGIYSHVHAKGVHWGLYKNGPWPMTTDEYT 629
Qy 228 -GNHDVASGSTIILAKQGEVWLQIFYSQNGLFYDPYWTDSFTGELI 275
Db 630 KGVLDQSSGSAIIDLFTENDQVWLQLPVASENGLYSSRY-VHSSFSGLV 677
RESULT 9
CA1A MOUSE
ID CA1A MOUSE STANDARD; PRT; 680 AA.
AC 005306;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN Name=Col10a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=93143676; PubMed=8424763;
RA Elima K., Berola I., Kovati R., Metsaeranta M., Garofalo S., Perala M.,
RA de Crombrughe B., Vuorio E.;
RT "The mouse collagen X gene: complete nucleotide sequence, exon
RT structure and expression pattern.";
RL Biochem. J. 289:247-253 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=93238750; PubMed=8477738;
RA Kong R.Y.C., Kwan K.M., Lau E.T., Thomas J.T., Boot-Handford R.P.,
RA Grant M.B., Cheah K.S.E.;
RT "Intron-exon structure, alternative use of promoter and expression of
RT the mouse collagen X gene, Col10a-1.";
RL Eur. J. Biochem. 213:99-111 (1993).
RN [3]
RP SEQUENCE OF 51-680 FROM N.A.
RC STRAIN=DBA/2J;
RX MEDLINE=92267014; PubMed=1587271;
RA Apic S.S., Seidin M.F., Hayashi M., Olsen B.R.;
RT "Cloning of the human and mouse type X collagen genes and mapping of
RT the mouse type X collagen gene to chromosome 10.";
RL Eur. J. Biochem. 206:217-224 (1992).
RN [4]
RP SEQUENCE OF 385-627 FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=92182017; PubMed=1543751;
RA Elima K., Metsaeranta M., Kallio J., Pernaalae M., Berola I.,
RA Garofalo S., de Crombrughe B., Vuorio E.;
RT "Specific hybridization probes for mouse alpha 2(I)X and alpha 1(X)
RT collagen mRNAs.";
RL Biochim. Biophys. Acta 1130:78-80 (1992).
RN [5]
RP FUNCTION: Type X collagen is a product of hyperthrophic
RP chondrocytes and has been localized to presumptive
RP mineralization zones of hyaline cartilage.
CC -1- SUBUNIT: Homotrimer.
CC -1- PM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -1- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@lab-sib.ch).
CC -----
CC EMBL; X67348; CAA47763.1; -
DR EMBL; X65121; CAA46237.1; -
DR EMBL; X63013; CAA44741.1; -
DR EMBL; Z21610; CAA79736.1; -
DR PIR; S31216; S31216.
DR HSSP; Q03692; IGR3.
DR MGJ; MG1:88445; COL10a1.
DR InterPro; IPR01073; C1q.
DR InterPro; IPR008161; C1q_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008983; TNF_1like.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 9.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 2.
DR PROSITE; PS01113; C1Q; 1.
KM Collagen; Connective tissue; Extracellular matrix; Hydroxylation;
KW Repeat; Signal.
FT SIGNAL; 1
FT CHAIN; 19
FT DOMAIN; 19
FT DOMAIN; 57
FT DOMAIN; 520
FT DOMAIN; 545
FT DOMAIN; 248
FT CONFLICT; 286
FT CONFLICT; 306
FT CONFLICT; 417
FT CONFLICT; 451
FT CONFLICT; 500
FT CONFLICT; 567
FT CONFLICT; 569
FT CONFLICT; 571
FT CONFLICT; 635
SQ SEQUENCE 680 AA; 66775 MW; FE984CA99A7F08E2 CRC64;

Query Match 32.3%; Score 511; DB 1; Length 680;
Best Local Similarity 36.4%; Pred. No. 4.8e-26;
Matches 114; Conservative 36; Mismatches 87; Indels 76; Gaps 6;

QY 38 LPFGQGGPRPGAGPGSGMGMGPRPGDGDHGDGRGDSGEGPP----- 84
DB 366 LVGAGAGPGARAGAPGGLDGTGYPSGPGNGPKMGNGLPQKSDPGVGSTPGLRGPPVG 425
QY 85 -----GRTGNRGKPGPKAGAIAGRAPGPGKGVNGTGGKHGTGPK----- 125
DB 426 PVGAKGVPGHNGGAPRGPGPIPTGRTGTPGVPVPGPPGSKDGNPGAPGAGIATKGL 485
QY 126 -----KGPKKKGPGPLPGP-----CSCGSG 146
DB 486 NGPTGPPGPPGPRGSGPGLPGPPGPPGQAVMPDGFIKAGQRPRLSGMPLVSNHG 545
QY 147 HT-----KAFSAVAVTKSYPRRLPIKPKDKILNBEGHVNASGKGVCGVGIYFTYDITL 203
DB 546 VTGMPVSAFTVILSKAIPAVGAPIPFDLILNRQGHDPGSGITCTCKIPGIYFSYHVH 605
QY 204 ANKLAIGLVHNGQYRIRTFDA-NTGNHNVASGSTITLAKOGDEVMILQIFYSEONGLFYD 262
DB 606 KGHVWGLYKNGCTPTMYTYDYBSKGYLDQASGSAIMELTENDQVWLQLPNAESNGLYSS 665
QY 263 PYWTDSLFTGFLI 275
DB 666 EY-VHSSFSGLV 677

RESULT 10
Q9Z1K4 PRELIMINARY; PRT; 295 AA.
AC Q9Z1K4;
DT 01-MAY-1999 (Tremblrel. 10; Created)
DT 01-MAY-1999 (Tremblrel. 10; Last sequence update)

```

```

DT 01-OCT-2003 (Tremblrel. 25; Last annotation update)
DE Collagen alpha 1 type X (Fragment).
GN Name=COL10A1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20310874; PubMed=10853827;
RA Marks S.A., Lundmark C., Christerson C., Wurtz T., Odgren P.R.,
RT Selfert M.F., Mackay C.A., Mason-Savas A., Popoff S.B.;
RT "Endochondral bone formation in toothless (osteopetrotic) rats:
RL failures of chondrocyte patterning and type X collagen expression.";
RL Int. J. Dev. Biol. 44:309-316(2000).
DR EMBL; AJ31848; CAA10518.1; -
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008983; TNF_1like.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 3.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen.
FT NON TER
SQ SEQUENCE 295 AA; 30012 MW; PF43B1548028813E CRC64;

Query Match 31.8%; Score 503; DB 2; Length 295;
Best Local Similarity 40.5%; Pred. No. 7.2e-26;
Matches 117; Conservative 32; Mismatches 90; Indels 50; Gaps 8;

QY 29 KGSFQVLCSLPFGPGPPGAPGSGMM--GMRGPPGKDGODHGDGRGDSGEGPPG 85
DB 12 KGNP-----GLPGPGGDDGVGAPGLRLVNGPTGKGVPGHNGGAPRGPGPIPTGRTGPIG 67
QY 86 RTGNRGKPGPKAGAIAGRAPG--DKGVNGTGGKHGTGPKKKGKGGPLPGP-- 140
DB 68 PGIIPGPPGSGDGGKPGAPGAGIVTKGLNGPA--GPRGRPPRGHTGPPGLPGPPG 124
QY 141 -----CSCGSGHK-----SAFSAVAVTKSYPRRLPI 167
DB 125 PGPPGPPSGAVIPDGFTRKSGQRPRLSGMPLVSNAGVTGMPVSAFTVILSKAIPAVGAPI 184
QY 168 KFDKILNBEGHVNASGKGVCGVGIYFTYDITLANKHLAIGLVHNGQYRIRTFDA-N 226
DB 185 PFDLILNRQGHDPGSGITCTCKIPGIYFSYHVHVGTHVWGLYKNGCTPTMYTYDYBS 244
QY 227 TGNHNVASGSTITLAKOGDEVMILQIFYSEONGLFYDYPWTDSLFTGFLI 275
DB 245 KGYLDQASGSAIMELTENDQVWLQLPNAESNGLYSSSEY-VHSSFSGLV 292

RESULT 11
ID CAIA BOVIN STANDARD; PRT; 674 AA.
AC P23206;
DT 01-NOV-1991 (Rel. 20; Created)
DT 01-NOV-1991 (Rel. 20; Last sequence update)
DT 05-JUN-2004 (Rel. 44; Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN Name=COL10A1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cartilage;

```

RX MEDLINE=9111313; PubMed=1703407;  
RA Thomas J.T., Kwan A.P.L., Grant M.B., Boot-Handford R.P.;  
RT "Isolation of cDNAs encoding the complete sequence of bovine type X  
RT collagen. Evidence for the condensed nature of mammalian type X  
RT collagen genes.";  
RL Biochem. J. 273:141-148(1991).  
CC -1- FUNCTION: Type X collagen is a product of hypertrophic  
CC chondrocytes and has been localized to presumptive  
CC mineralization zones of hyaline cartilage.  
CC -1- SUBUNIT: Homotrimer.  
CC -1- PM: Prolines at the third position of the tripeptide repeating  
CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
CC -1- SIMILARITY: Contains 1 C1q domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, X53556; CA37624.1; -.  
DR PIR, S13301; S13301.  
DR HSSP, Q03692; IGR3.  
DR InterPro: IPR001073; C1q.  
DR InterPro: IPR008161; C1q helix.  
DR InterPro: IPR008160; Collagen.  
DR InterPro: IPR008983; TNF-like.  
DR Pfam; PF00386; C1q; 1.  
DR Pfam; PF01391; Collagen; 9.  
DR PRINTS; PRO0007; COMPLEMENTC1Q.  
DR Prodom; PD000007; C1q helix; 2.  
DR PROSITE; PS01113; C1Q; 1.  
KW Collagen; Connective tissue; Extracellular matrix; Glycoprotein;  
KW Hydroxylation; Repeat; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 674 Collagen alpha 1(X) chain.  
FT DOMAIN 19 56 Nonhelical region (NC2).  
FT DOMAIN 57 519 Triple-helical region.  
FT DOMAIN 520 674 Nonhelical region (NC1).  
FT DOMAIN 539 674 C1q.  
FT DISULFID 194 197 By similarity.  
FT MOD\_RES 460 460 Hydroxyproline (By similarity).  
FT MOD\_RES 463 463 Hydroxyproline (By similarity).  
FT CARBOHYD 611 611 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 674 AA; 65546 MW; CD4CA73A03E004CA CRC64;

Query Match 31.8%; Score 502.5; DB 1; Length 674;  
Best Local Similarity 41.8%; Pred. No. 1.7e-25;  
Matches 117; Conservative 28; Mismatches 94; Indels 41; Gaps 8;

QY 29 KGSFOLVCSLPFGQGP---GPPGAPGSGMGMGPPGKDGODGHDGDRGDSGEEGPPG 85  
DB 400 KGNP---GLPGRKGDPGIGAGSPGLPGVPAGAKGVPGHNGEAGPRGVPGLPSTRPPIG 455  
QY 86 RTGNRGKPRGPKGAKGAIGRAPRGPKG--VNGTPGKHGTGPKGPKKKGEPGLPGP--- 140  
DB 456 PPGIPGPPGSKD---VGTPEPBPAGIAYKGLNGLPGPBPBPBPKNAGBPGLPGPPGP 512  
QY 141 -----CSCGSGHT---KSASFVAVTKSPRRRLPIKFDKILMNE 176  
DB 513 PGPFGVALPDEDFYKAGRPVSNQGVTMGPVSAFTVILSKAYPAIGTPIPPDKILYNK 572  
QY 177 GGHYNASSGFVGVGPIYFTYDITLANKHLAIGLVHNGQYRIKTPDAN--GNHVDVAG 235  
DB 573 QOHYDPRGTGFTCRIGIYFYSYHINVKGTAWGLYKNGTPVWYTYDEVYKGYLDASG 632  
QY 236 STILALKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLI 275  
DB 633 SAVIDLTENDQVWLQLPNAGSNGL--YSPRYVHSSFGFLV 671

RESULT 12  
Q9N178 PRELIMINARY; PRT; 675 AA.  
AC Q9N178;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Type X collagen.  
GN Name=COL10A1;  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21015405; PubMed=11130976;  
RA Nielsen V.H., Bendixen C., Ambjerg J., Sorensen C.M., Jensen H.B.,  
RA Shukri N.M., Thomson B.;  
RT "Abnormal growth plate function in pigs carrying a dominant mutation  
RT in type X collagen.";  
RL Mamm. Genome 11:1087-1092(2000).  
DR EMBL, AF222861; AAF37271.1; -.  
DR HSSP, Q03692; IGR3.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro: IPR001073; C1q.  
DR InterPro: IPR008161; C1q helix.  
DR InterPro: IPR008160; Collagen.  
DR InterPro: IPR008983; TNF-like.  
DR Pfam; PF00386; C1q; 1.  
DR Pfam; PF01391; Collagen; 8.  
DR PRINTS; PRO0007; COMPLEMENTC1Q.  
DR Prodom; PD000007; C1q helix; 3.  
DR SMART; SM00110; C1Q; 1.  
DR PROSITE; PS01113; C1Q; 1.  
KW Collagen.  
SQ SEQUENCE 675 AA; 65447 MW; 26397B10310383F9 CRC64;

Query Match 31.8%; Score 502.5; DB 2; Length 675;  
Best Local Similarity 42.5%; Pred. No. 1.7e-25;  
Matches 119; Conservative 23; Mismatches 97; Indels 41; Gaps 8;

QY 29 KGSFOLVCSLPFGQGP---GPPGAPGSGMGMGPPGKDGODGHDGDRGDSGEEGPPG 85  
DB 401 KGNP---GLPGRKGDPGIGAGSPGLPGVPAGAKGVPGHNGEAGPRGVPGLPSTRPPIG 456  
QY 86 RTGNRGKPRGPKGAKGAIGRAPRGPKG--VNGTPGKHGTGPKGPKKKGEPGLPGP--- 140  
DB 457 PPGIPGPPGSKD---GNPBPBPAGIATKGLNGLPGPBPBPBPKNAGBPGLPGPPGP 513  
QY 141 -----CSCGSGHT---KSASFVAVTKSPRRRLPIKFDKILMNE 176  
DB 514 PGPFGVALPDEDFYKAGRPVSNQGVTMGPVSAFTVILSKAYPAIGTPIPPDKILYNK 573  
QY 177 GGHYNASSGFVGVGPIYFTYDITLANKHLAIGLVHNGQYRIKTPDAN--GNHVDVAG 235  
DB 574 QOHYDPRGTGFTCRIGIYFYSYHINVKGTAWGLYKNGTPVWYTYDEVYKGYLDASG 633  
QY 236 STILALKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLI 275  
DB 634 SAVIDLTENDQVWLQLPNAGSNGLYSSEY--VHSSFGFLV 672

RESULT 13  
CA1A CHICK STANDARD; PRT; 674 AA.  
AC P08125;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 05-JUN-2004 (Rel. 44, Last annotation update)  
DE Collagen alpha 1(X) chain precursor.  
GN Name=COL10A1;  
OS Gallus gallus (Chicken).





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 25, 2004, 08:12:38 / Search time 123.56 Seconds

(without alignments)  
152.967 Million cell updates/sec

Title: US-10-621-787-2

Perfect score: 1582  
Sequence: 1 MIPWVLAALPCADPLG.....TDSUPFLIYADDDNEV 285Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/PCrus\_COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	1582	100.0	285	4	US-09-552-204A-2
2	1487	94.0	294	3	US-09-188-930-294
3	1487	94.0	294	4	US-09-312-283C-294
4	1484	93.8	285	4	US-09-312-283C-382
5	981	62.0	198	3	US-09-186-930-138
6	981	62.0	198	4	US-09-312-283C-138
7	553.5	32.5	120	4	US-09-552-204A-12
8	513.5	32.5	162	3	US-09-227-357-436
9	431.5	27.3	244	2	US-08-463-911-7
10	431.5	27.3	244	3	US-09-140-804-3
11	431.5	27.3	244	4	US-09-336-536-20
12	431.5	27.3	244	4	US-09-530-423-1
13	431.5	27.3	244	4	US-09-686-838B-3
14	431.5	27.3	244	4	US-09-911-176B-48
15	431.5	27.3	244	4	US-09-552-225A-3
16	431.5	27.3	244	4	US-09-619-740-51
17	431.5	27.3	244	4	US-09-776-976-6
18	431.5	27.3	244	4	US-09-909-547-6
19	431.5	27.3	244	4	US-09-552-204A-3
20	425	26.9	247	2	US-08-463-911-2
21	425	26.9	247	4	US-09-776-976-4
22	425	26.9	247	4	US-09-909-547-4
23	424.5	26.8	244	4	US-09-568-852B-6
24	420.5	26.6	231	4	US-09-530-423-2
25	420	26.5	423	1	US-08-383-744-2
26	420	26.5	423	2	US-08-999-336-2
27	420	26.5	423	5	PCT-US96-01427-2

28	417	26.4	247	3	US-09-140-804-8	Sequence 8, Appli
29	417	26.4	247	3	US-09-118-408-3	Sequence 3, Appli
30	417	26.4	247	4	US-09-506-885-3	Sequence 3, Appli
31	417	26.4	247	4	US-09-686-838B-8	Sequence 8, Appli
32	417	26.4	247	4	US-09-911-176B-3	Sequence 3, Appli
33	417	26.4	247	4	US-09-619-740-3	Sequence 3, Appli
34	417	26.4	247	4	US-09-506-882-3	Sequence 3, Appli
35	415	26.2	243	3	US-09-188-930-295	Sequence 295, App
36	415	26.2	243	4	US-09-312-283C-295	Sequence 295, App
37	410	25.9	247	4	US-09-776-976-2	Sequence 2, Appli
38	410	25.9	247	4	US-09-909-547-2	Sequence 2, Appli
39	408	25.8	243	4	US-09-336-536-11	Sequence 11, Appli
40	403.5	25.5	228	4	US-09-336-536-10	Sequence 10, Appli
41	403	25.5	243	3	US-09-140-804-2	Sequence 2, Appli
42	403	25.5	243	4	US-09-336-536-2	Sequence 2, Appli
43	403	25.5	243	4	US-09-686-838B-2	Sequence 2, Appli
44	403	25.5	243	4	US-09-866-028-12	Sequence 42, Appli
45	403	25.5	243	4	US-10-140-002-362	Sequence 42, Appli
46	403	25.5	243	4	US-09-944-457-42	Sequence 42, Appli
47	398.5	25.2	228	4	US-09-336-536-4	Sequence 4, Appli
48	393	24.8	246	2	US-08-463-911-4	Sequence 4, Appli
49	366	23.1	245	4	US-09-552-225A-4	Sequence 4, Appli
50	366	23.1	245	4	US-09-552-204A-4	Sequence 4, Appli
51	355.5	22.5	258	4	US-09-976-594-815	Sequence 815, App
52	349.5	22.5	245	4	US-09-800-729-116	Sequence 116, App
53	348.5	22.0	245	3	US-09-140-804-4	Sequence 4, Appli
54	348.5	22.0	245	4	US-09-686-838B-4	Sequence 4, Appli
55	348.5	22.0	245	4	US-09-911-176B-49	Sequence 49, Appli
56	348.5	22.0	245	4	US-09-619-740-52	Sequence 52, Appli
57	348.5	22.0	245	4	US-09-800-729-85	Sequence 85, Appli
58	348.5	22.0	245	4	US-09-800-729-115	Sequence 115, App
59	348.5	22.0	245	4	US-09-800-729-118	Sequence 118, App
60	348.5	22.0	245	4	US-09-800-729-119	Sequence 119, App
61	345.5	21.8	245	4	US-09-311-021-104	Sequence 104, App
62	324	20.5	229	4	US-09-800-729-117	Sequence 117, App
63	311.5	19.7	1057	3	US-08-931-820-1	Sequence 1, Appli
64	311.5	19.7	1461	4	US-09-585-887-9	Sequence 9, Appli
65	311.5	19.7	1461	4	US-09-289-578-9	Sequence 9, Appli
66	311.5	19.7	1464	4	US-09-331-377C-21	Sequence 21, Appli
67	306	19.3	215	3	US-09-140-804-5	Sequence 5, Appli
68	306	19.3	215	4	US-09-686-838B-5	Sequence 5, Appli
69	306	19.3	215	4	US-09-911-176B-50	Sequence 50, Appli
70	306	19.3	215	4	US-09-619-740-53	Sequence 53, Appli
71	305.5	19.3	246	4	US-09-552-225A-12	Sequence 12, Appli
72	305	19.3	1341	3	US-08-963-825-18	Sequence 18, Appli
73	305	19.3	1341	3	US-09-500-811-18	Sequence 18, Appli
74	305	19.3	1341	3	US-09-570-573-18	Sequence 18, Appli
75	305	19.3	1341	3	US-09-548-608-18	Sequence 18, Appli
76	304	19.2	246	4	US-09-552-225A-2	Sequence 2, Appli
77	304	19.2	319	4	US-10-012-605C-2	Sequence 2, Appli
78	303	19.2	492	4	US-08-468-996-12	Sequence 12, Appli
79	302	19.1	310	3	US-09-219-849-17	Sequence 47, Appli
80	302	19.1	595	3	US-09-219-849-18	Sequence 48, Appli
81	302	19.1	595	3	US-09-219-849-50	Sequence 50, Appli
82	302	19.1	822	3	US-09-219-849-49	Sequence 49, Appli
83	293	18.5	1064	3	US-08-642-255-62	Sequence 62, Appli
84	288.5	18.2	1078	3	US-08-963-825-21	Sequence 21, Appli
85	288.5	18.2	1078	3	US-09-500-811-21	Sequence 21, Appli
86	288.5	18.2	1078	3	US-09-570-573-21	Sequence 21, Appli
87	288.5	18.2	1078	3	US-09-548-608-21	Sequence 21, Appli
88	288	18.2	330	1	US-08-642-255-12	Sequence 32, Appli
89	288	18.2	408	1	US-07-609-716-65	Sequence 65, Appli
90	288	18.2	408	3	US-08-475-411A-65	Sequence 65, Appli
91	288	18.2	408	3	US-08-478-029A-65	Sequence 65, Appli
92	286.5	18.1	1057	3	US-08-931-820-4	Sequence 4, Appli
93	286	18.1	1418	3	US-08-963-825-20	Sequence 20, Appli
94	286	18.1	1418	3	US-09-010-999-1	Sequence 1, Appli
95	286	18.1	1418	3	US-09-500-811-20	Sequence 20, Appli
96	286	18.1	1418	3	US-09-570-573-20	Sequence 20, Appli
97	286	18.1	1418	3	US-09-548-608-20	Sequence 20, Appli
98	285.5	18.0	252	1	US-08-642-255-61	Sequence 61, Appli
99	285	18.0	357	1	US-07-609-716-66	Sequence 66, Appli
100	285	18.0	357	1	US-08-642-255-33	Sequence 33, Appli

## ALIGNMENTS

RESULT 1  
 US-09-552-204A-2  
 ; Sequence 2, Application US/09552204A  
 ; Patent No. 6620909  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fiddington, Christopher S.  
 ; APPLICANT: Bishop, Paul D.  
 ; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACR2P  
 ; FILE REFERENCE: 99-08  
 ; CURRENT APPLICATION NUMBER: US/09/552,204A  
 ; PRIOR FILING DATE: 2000-04-19  
 ; PRIOR APPLICATION NUMBER: 60/130,207  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 285  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-552-204A-2

Query Match 100.0%; Score 1582; DB 4; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-132;  
 Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPQPPGAPGSPGMMGRM 60  
 DB 1 MIPWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPQPPGAPGSPGMMGRM 60

QY 61 GPPKDGQDQDHDGRGSGSEGGPPGRTGNRKGKPKKAKAIGAGRGPGVNGTGK 120  
 DB 61 GPPKDGQDQDHDGRGSGSEGGPPGRTGNRKGKPKKAKAIGAGRGPGVNGTGK 120

QY 121 GTPKPKPKKKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 180  
 DB 121 GTPKPKPKKKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 180

QY 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDV 240  
 DB 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDV 240

QY 241 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285  
 DB 241 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285

RESULT 2  
 US-09-188-930-294  
 ; Sequence 294, Application US/09188930A  
 ; Patent No. 6150502  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Watson, James D.  
 ; APPLICANT: Strachan, Lorna  
 ; APPLICANT: Sleeman, Matthew  
 ; APPLICANT: Onrust, Rene  
 ; APPLICANT: Murison, James Greg  
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
 ; TITLE OF INVENTION: and Methods For Their Use  
 ; FILE REFERENCE: 11000,1011c1  
 ; CURRENT APPLICATION NUMBER: US/09/188,930A  
 ; CURRENT FILING DATE: 1998-11-09  
 ; NUMBER OF SEQ ID NOS: 348  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 294  
 ; LENGTH: 294  
 ; TYPE: PRT  
 ; ORGANISM: Rat  
 US-09-188-930-294

Query Match 94.0%; Score 1487; DB 3; Length 294;  
 Best Local Similarity 94.0%; Pred. No. 6,4e-124;  
 Matches 268; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 MIPWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPQPPGAPGSPGMMGRM 60  
 DB 10 MISWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPQPPGAPGSPGMMGRM 69

QY 61 GPPKDGQDQDHDGRGSGSEGGPPGRTGNRKGKPKKAKAIGAGRGPGVNGTGK 120  
 DB 70 GPPKDGQDQDHDGRGSGSEGGPPGRTGNRKGKPKKAKAIGAGRGPGVNGTGK 129

QY 121 GTPKPKPKKKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 180  
 DB 130 GTPKPKPKKKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 189

QY 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDV 240  
 DB 190 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDV 249

QY 241 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285  
 DB 250 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 294

RESULT 3  
 US-09-312-283C-294  
 ; Sequence 294, Application US/09312283C  
 ; Patent No. 6573095  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Watson, James D.  
 ; APPLICANT: Strachan, Lorna  
 ; APPLICANT: Sleeman, Matthew  
 ; APPLICANT: Onrust, Rene  
 ; APPLICANT: Murison, James G.  
 ; APPLICANT: Kumble, Krishanand D.  
 ; TITLE OF INVENTION: Compositions Isolated from Skin Cells  
 ; TITLE OF INVENTION: and Methods for Their Use  
 ; FILE REFERENCE: 11000,1011c2  
 ; CURRENT APPLICATION NUMBER: US/09/312,283C  
 ; CURRENT FILING DATE: 1999-05-14  
 ; NUMBER OF SEQ ID NOS: 425  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 294  
 ; LENGTH: 294  
 ; TYPE: PRT  
 ; ORGANISM: Mouse  
 US-09-312-283C-294

Query Match 94.0%; Score 1487; DB 4; Length 294;  
 Best Local Similarity 94.0%; Pred. No. 6,4e-124;  
 Matches 268; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 MIPWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPQPPGAPGSPGMMGRM 60  
 DB 10 MISWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPQPPGAPGSPGMMGRM 69

QY 61 GPPKDGQDQDHDGRGSGSEGGPPGRTGNRKGKPKKAKAIGAGRGPGVNGTGK 120  
 DB 70 GPPKDGQDQDHDGRGSGSEGGPPGRTGNRKGKPKKAKAIGAGRGPGVNGTGK 129

QY 121 GTPKPKPKKKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 180  
 DB 130 GTPKPKPKKKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 189

QY 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDV 240  
 DB 190 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDV 249

QY 241 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285  
 DB 250 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 294



RESULT 4  
US-09-312-283C-382  
Sequence 382, Application US/09312283C  
Patent No. 6573095  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions Isolated from Skin Cells  
FILE REFERENCE: 11000.1011c2  
CURRENT APPLICATION NUMBER: US/09/312,283C  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 425  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 382  
LENGTH: 285  
TYPE: PRT  
ORGANISM: Mouse  
US-09-312-283C-382

Query Match 93.8%; Score 1484; DB 4; Length 285;  
Best Local Similarity 93.7%; Pred. No. 1,1e-123;  
Matches 267; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCADPPLGAFARDFRKSPQLVCSLPFGQPPGPGAGPAGSGMGRM 60  
DB 1 MISWMLLACALPCADPPLGAFARDFRKSPQLVCSLPFGQPPGPGAGSGMGRM 60  
QY 61 GFGPKDQDGDHGRGDSGSEGGPPGRTGNRKGPKKAGAGIAGRAGRGPGVNGTGGKH 120  
DB 61 GFGPKDQDGDHGRGDSGSEGGPPGRTGNRKGPKKAGAGIAGRAGRGPGVNGTGGKH 120  
QY 121 GTPGKKGPKKGGKGGPGLPGPCSSCGSGHTKSAFSAVAVTKSYRERLPIKFDKILNNEGHHY 180  
DB 121 GTPGKKGPKKGGKGGPGLPGPCSSCGSGHTKSAFSAVAVTKSYRERLPIKFDKILNNEGHHY 180  
QY 181 NASGKRFVCVCGIYFTYDITLANKHLAIGLVHNGVORITPPANTGNHVAAGSTLLA 240  
DB 181 NASGKRFVCVCGIYFTYDITLANKHLAIGLVHNGVORITPPANTGNHVAAGSTLLA 240  
QY 241 LKQDEVWLQIFYSQNGLFTDPYMTDSLFTGFLIYADQDDPNRY 285  
DB 241 LKQDEVWLQIFYSQNGLFTDPYMTDSLFTGFLIYADQDDPNRY 285

RESULT 5  
US-09-188-930-138  
Sequence 138, Application US/09188930A  
Patent No. 6150502  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James G.  
TITLE OF INVENTION: Compositions Isolated from Skin Cells  
FILE REFERENCE: 11000.1011c1  
CURRENT APPLICATION NUMBER: US/09/188,930A  
CURRENT FILING DATE: 1998-11-09  
NUMBER OF SEQ ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 138  
LENGTH: 198  
TYPE: PRT  
ORGANISM: Rat  
US-09-188-930-138

Query Match 62.0%; Score 981; DB 3; Length 198;  
Best Local Similarity 92.6%; Pred. No. 2.6e-79;  
Matches 175; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCADPPLGAFARDFRKSPQLVCSLPFGQPPGPGAGPAGSGMGRM 60  
DB 10 MISWMLLACALPCADPPLGAFARDFRKSPQLVCSLPFGQPPGPGAGSGMGRM 69  
QY 61 GFGPKDQDGDHGRGDSGSEGGPPGRTGNRKGPKKAGAGIAGRAGRGPGVNGTGGKH 120  
DB 70 GFGPKDQDGDHGRGDSGSEGGPPGRTGNRKGPKKAGAGIAGRAGRGPGVNGTGGKH 129  
QY 121 GTPGKKGPKKGGKGGPGLPGPCSSCGSGHTKSAFSAVAVTKSYRERLPIKFDKILNNEGHHY 180  
DB 130 GTPGKKGPKKGGKGGPGLPGPCSSCGSGHTKSAFSAVAVTKSYRERLPIKFDKILNNEGHHY 189  
QY 181 NASGKRFVC 189  
DB 190 NASGKRFVC 198

RESULT 6  
US-09-312-283C-138  
Sequence 138, Application US/09312283C  
Patent No. 6573095  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions Isolated from Skin Cells  
FILE REFERENCE: 11000.1011c2  
CURRENT APPLICATION NUMBER: US/09/312,283C  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 425  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 138  
LENGTH: 198  
TYPE: PRT  
ORGANISM: Mouse  
US-09-312-283C-138

Query Match 62.0%; Score 981; DB 4; Length 198;  
Best Local Similarity 92.6%; Pred. No. 2.6e-79;  
Matches 175; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCADPPLGAFARDFRKSPQLVCSLPFGQPPGPGAGPAGSGMGRM 60  
DB 10 MISWMLLACALPCADPPLGAFARDFRKSPQLVCSLPFGQPPGPGAGSGMGRM 69  
QY 61 GFGPKDQDGDHGRGDSGSEGGPPGRTGNRKGPKKAGAGIAGRAGRGPGVNGTGGKH 120  
DB 70 GFGPKDQDGDHGRGDSGSEGGPPGRTGNRKGPKKAGAGIAGRAGRGPGVNGTGGKH 129  
QY 121 GTPGKKGPKKGGKGGPGLPGPCSSCGSGHTKSAFSAVAVTKSYRERLPIKFDKILNNEGHHY 180  
DB 130 GTPGKKGPKKGGKGGPGLPGPCSSCGSGHTKSAFSAVAVTKSYRERLPIKFDKILNNEGHHY 189  
QY 181 NASGKRFVC 189  
DB 190 NASGKRFVC 198

RESULT 7  
US-09-552-204A-12  
Sequence 12, Application US/09552204A  
Patent No. 6620909  
GENERAL INFORMATION:  
APPLICANT: Piddington, Christopher S.

APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACRP2  
FILE REFERENCE: 99-08  
CURRENT APPLICATION NUMBER: US/09/552,204A  
CURRENT FILING DATE: 2000-04-19  
PRIOR APPLICATION NUMBER: 60/130,207  
PRIOR FILING DATE: 1999-04-20  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 120  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)-(120)  
OTHER INFORMATION: Each Xaa is independently any amino acid.  
US-09-552-204A-12

Query Match 35.2%; Score 556.5; DB 4; Length 120;  
Best Local Similarity 88.4%; Pred. No. 5.6e-42;  
Matches 107; Conservative 2; Mismatches 9; Indels 3; Gaps 2;

QY 167 IKPKILMNEGGHYNASGKFCVCGVPGIYTYDITL--ANKHLAIGLVNNGQYRIRTFD 224  
DB 1 IKPKILMNEGGHYNASGKFCVCGVPGI--XIXLIPMTLRXANKKXKXIGLVNNGQYRIRTFD 59  
QY 225 ANTGNHDVAGSSTILALKQGDVWMLQIIFYSEONGLFYDPYWTDSLFTGFLIYADODDPE 284  
DB 60 ANTGNHDVAGSSTILALKQGDVWMLQIIFYSEONGLFYDPYWTDSLFTGFLIYADODDPE 119  
QY 285 V 285  
DB 120 V 120

RESULT 8  
US-09-227-357-436  
Sequence 436, Application US/09227357  
Patent No. 6342581  
GENERAL INFORMATION:  
APPLICANT: Fischer et al.  
TITLE OF INVENTION: 123 Human Secreted Proteins  
FILE REFERENCE: P2010P1  
CURRENT APPLICATION NUMBER: US/09/227,357  
CURRENT FILING DATE: 1999-01-08  
EARLIER APPLICATION NUMBER: PCT/US98/13684  
EARLIER FILING DATE: 1998-07-07  
EARLIER APPLICATION NUMBER: 60/051,926  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,793  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,925  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,929  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,803  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,732  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,931  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,932  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,916  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,930  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,918  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,920  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,733

EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,795  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,919  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,928  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/055,722  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,723  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,948  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,949  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,953  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,950  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,947  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,964  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/056,360  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,684  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,984  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,954  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/058,785  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,664  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,660  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,661  
EARLIER FILING DATE: 1997-09-12  
NUMBER OF SEQ ID NOS: 672  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 436  
LENGTH: 162  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (1)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (33)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (48)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-227-357-436

Query Match 32.5%; Score 513.5; DB 3; Length 162;  
Best Local Similarity 65.1%; Pred. No. 5.3e-38;  
Matches 99; Conservative 17; Mismatches 27; Indels 9; Gaps 2;

QY 134 BPGLPBCSGSGHGTSAFSAVAATKSYPRRLPIKPKILMNEGG-----HYNASGKFCV 188  
DB 4 DPGLPBCSGSIVLKAFAFVGITTSYPRRLPIKPKILMNEGG-----HYNASGKFCV 59  
QY 189 CGVPGIYTYDITLANKHLAIGLVNNGQYRIRTFDANTGNHDVAGSSTILALKQGDVW 248  
DB 60 VLSGDIYFSDITLANKHLAIGLVNNGQYRIRTFDANTGNHDVAGSSTIVLYLPEDVW 119  
QY 249 LQIFYSEONGLFYDPYWTDSLFTGFLIYADOD 280

Db 120 LEIFFTDQNGLFSDPGMADSLFSGFLLYVTD 151

RESULT 9

US-08-463-911-7

Sequence 7, Application US/08463911

Patent No. 5869330

GENERAL INFORMATION:

APPLICANT: Scherer, Philipp E.

APPLICANT: Lodish, Harvey F.

TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED

TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Millitia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,911

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: WHI95-05

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 244 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-463-911-7

Query Match 27.3%; Score 431.5; DB 2; Length 244;

Best Local Similarity 40.2%; Pred. No. 1.6e-30;

Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPQLVCSLPG-----PGGP-----PGPPGAPGSPGMMGMPGPKDQGDHDDRDSDG 79

Db 5 GAVLLALALPGHDDQETTTQGPVLLPLPKGA--CTGMA--GIPGHGHN-----50

Qy 80 EEGPPGRTGNRGKGPCKAGAIAGAPRGPGKGVNTPGKHGTGPKKGPCKGEPGLPG 139

Db 51 --GAPGRDGRDTGEGKEKDPGLIGPKDIGFTGVPGAGPRGPGIQRKEPGE--106

Qy 140 PCSCSGHTSASFVAATYKSPRRLPIKFKDKILMNEGSHYNASSGKFCVCPGIYFTY 199

Db 107 ----GAVYRSAPFVGLETYYTINMPIRFTKIFYNOQNHDSGTGKFKCNIPGLYYFAY 162

Qy 200 DITLANHGLAIGLVHNGQYRIRTPDA--NTGNHDVASSSTIALXQGBEVLQIF--YSEQN 257

Db 163 HITVYMDVKVSLFKKOKAMLFITYDOENNVDAQSSVLLHLBVGQVWLQVYGEGERN 222

Qy 258 GLFYDPYWTDSLFTGFLIYAD 278

Db 223 GLYADND--NDSTFTGFLIYHD 242

RESULT 10

US-09-140-804-3

Sequence 3, Application US/09140804

Patent No. 6197930

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Humes, Jacqueline M.

TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS

FILE REFERENCE: 97-49

CURRENT APPLICATION NUMBER: US/09/140,804

CURRENT FILING DATE: 1998-08-26

EARLIER APPLICATION NUMBER: 60/056,983

EARLIER FILING DATE: 1997-08-26

NUMBER OF SEQ ID NOS: 47

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 3

LENGTH: 244

TYPE: PRT

ORGANISM: Homo sapiens

US-09-140-804-3

Query Match 27.3%; Score 431.5; DB 3; Length 244;

Best Local Similarity 40.2%; Pred. No. 1.6e-30;

Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPQLVCSLPG-----PGGP-----PGPPGAPGSPGMMGMPGPKDQGDHDDRDSDG 79

Db 5 GAVLLALALPGHDDQETTTQGPVLLPLPKGA--CTGMA--GIPGHGHN-----50

Qy 80 EEGPPGRTGNRGKGPCKAGAIAGAPRGPGKGVNTPGKHGTGPKKGPCKGEPGLPG 139

Db 51 --GAPGRDGRDTGEGKEKDPGLIGPKDIGFTGVPGAGPRGPGIQRKEPGE--106

Qy 140 PCSCSGHTSASFVAATYKSPRRLPIKFKDKILMNEGSHYNASSGKFCVCPGIYFTY 199

Db 107 ----GAVYRSAPFVGLETYYTINMPIRFTKIFYNOQNHDSGTGKFKCNIPGLYYFAY 162

Qy 200 DITLANHGLAIGLVHNGQYRIRTPDA--NTGNHDVASSSTIALXQGBEVLQIF--YSEQN 257

Db 163 HITVYMDVKVSLFKKOKAMLFITYDOENNVDAQSSVLLHLBVGQVWLQVYGEGERN 222

Qy 258 GLFYDPYWTDSLFTGFLIYAD 278

Db 223 GLYADND--NDSTFTGFLIYHD 242

RESULT 11

US-09-336-536-20

Sequence 20, Application US/09336536

Patent No. 6406884

GENERAL INFORMATION:

APPLICANT: Leidy, K.

APPLICANT: McKay, C.

APPLICANT: Boesone, S.

TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF

FILE REFERENCE: 7853-144

CURRENT APPLICATION NUMBER: US/09/336,536

CURRENT FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 75

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 20

LENGTH: 244

TYPE: PRT

ORGANISM: Homo sapiens

US-09-336-536-20

Query Match 27.3%; Score 431.5; DB 4; Length 244;

Best Local Similarity 40.2%; Pred. No. 1.6e-30;

Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPQLVCSLPG-----PGGP-----PGPPGAPGSPGMMGMPGPKDQGDHDDRDSDG 79

Db 5 GAVLLALALPGHDDQETTTQGPVLLPLPKGA--CTGMA--GIPGHGHN-----50

Qy 80 EEGPPGRTGNRGKGPCKAGAIAGAPRGPGKGVNTPGKHGTGPKKGPCKGEPGLPG 139

Db 51 --GAPGRDGRDTGEGKEKDPGLIGPKDIGFTGVPGAGPRGPGIQRKEPGE--106

	Query Match	27.3%	Score 431.5;	DB 4;	length 244;
	Best Local Similarity	40.2%;	Pred. No. 1,6e-30;		
	Matches	105;	Conservative	33;	Mismatches 88; Indels 35; Gaps 9;
Qy	30	GSPQLVCSLPG-----PQGP----	PGPPGAPGPSGMNGRMPFGPKDGDGHDGRDSDG	79	
Db	5	GAVALLLALPHDDOETTTQQGPVLLPLRKGA--CTGWMA--GIPIHPHN-----		50	
Qy	80	EEGPPGRTGNKGPPEPKKAGALIGRAGRGPKNVGVTGKATPEPKKPRPKGKGEPGLPG	139		
Db	51	--GAPGRDRGTPEKEKEKDPLGIGRKDIGEIVGVPAGSGRPFPGLIQRKSGPGE--	106		
Qy	140	PCSCGSGETKAFSAVAVTKSYPRERLPIKFDKILMNEGHYNASGGKVCGVPGIYYTTY	199		
Db	107	----GAYYRAAFSAVGLTEYYTIPNMPIRFTKIFLNQONHVDGASGKHCHINIPGLYYFAY	162		
Qy	200	DITLANIKHALGLVINGOYRIRTFDA--NTGNHDVASGSTIALKOGDEVMLOIF--YSEON	257		
Db	163	HITYVMKDVKSILFKDKAMLFETYDQYENNVDAQSSGVLLHLBVGDDQVMIVQVYBGBERN	222		
Qy	258	GLPTDPYWTDSLFTGFLFYAD	278		
Db	223	GLYADND--NDSTFGFFLLYHD	242		

```

; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: US 60/056,983
; PRIOR FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 244
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-686-838B-3

Query Match 27.3%; Score 431.5; DB 4; Length 244;
Beat Local Similarity 40.2%; Pred. No. 1.6e-30;
Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

QY 30 GSPQVLSLG-----PQGP-----PGPPGAPGSGMMGMGFPKXGQGDGHDGRDSDG 79
Db 5 GAVLLALPQHDDQETTTQGGVLLPLPKGH--CTGMA--GLPGHGHN----- 50
QY 80 EEGPPGRTGNKGPCKGPKAGAIGRAPRGPKVNGTGPKNHTGPKKKGPKKGPGLPG 139
Db 51 --GAPGDRGDRGTGEGKEGKDPGLIGPKGIDGTVGPAEGPRGPGFQIGRKGPGB-- 106
QY 140 PCSGSGHTGSAFSAVATKSYPRRRLIKPKDKILMEGCHYMASGKFCVGPRIYYFTY 199
Db 107 ---GAVYSAFSAFVSGLETYYTLIPNMELRFTKLFYNNQNHDSGTGPKCNIPELYFYAY 162
QY 200 DITLANGLAIGLVHNCQYRIREFDA-NTGNHDVASGTTILAKOGDEWLVQIF-VSEON 257
Db 163 HITVYMDVAKSLPKKKOKAMLFYTDQYQENNVDAQSSVLLHLEVGQVWLQVYGEERN 222
QY 258 GLFYDPYWTDSLFTGFLIYAD 278
Db 223 GLYADND-NDSTFTGFLIYAD 242

RESULT 14
US-09-911-176B-48
; Sequence 48, Application US/09911176B
; Patent No. 6518403
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ANTIBODIES THAT BIND AN
; TITLE OR INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG
; FILE REFERENCE: 97-30D1
; CURRENT APPLICATION NUMBER: US/09/911.176B
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/118,408
; PRIOR FILING DATE: 1998-07-17
; PRIOR APPLICATION NUMBER: 60/053,154
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 48
; LENGTH: 244
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-911-176B-48

Query Match 27.3%; Score 431.5; DB 4; Length 244;
Beat Local Similarity 40.2%; Pred. No. 1.6e-30;
Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

QY 30 GSPQVLSLG-----PQGP-----PGPPGAPGSGMMGMGFPKXGQGDGHDGRDSDG 79
Db 5 GAVLLALPQHDDQETTTQGGVLLPLPKGH--CTGMA--GLPGHGHN----- 50
QY 80 EEGPPGRTGNKGPCKGPKAGAIGRAPRGPKVNGTGPKNHTGPKKKGPKKGPGLPG 139
Db 51 --GAPGDRGDRGTGEGKEGKDPGLIGPKGIDGTVGPAEGPRGPGFQIGRKGPGB-- 106
QY 140 PCSGSGHTGSAFSAVATKSYPRRRLIKPKDKILMEGCHYMASGKFCVGPRIYYFTY 199
Db 107 ---GAVYSAFSAFVSGLETYYTLIPNMELRFTKLFYNNQNHDSGTGPKCNIPELYFYAY 162
QY 200 DITLANGLAIGLVHNCQYRIREFDA-NTGNHDVASGTTILAKOGDEWLVQIF-VSEON 257
Db 163 HITVYMDVAKSLPKKKOKAMLFYTDQYQENNVDAQSSVLLHLEVGQVWLQVYGEERN 222
QY 258 GLFYDPYWTDSLFTGFLIYAD 278
Db 223 GLYADND-NDSTFTGFLIYAD 242

```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:38:38 / Search time 501.456 Seconds  
(Without alignments)  
204.102 Million cell updates/sec

Title: US-10-621-787-2  
Perfect score: 1582  
Sequence: 1 MIRVWLLACALPCADPLG.....TDSLPTGFLYADDDNEV 285

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1595201 seqs, 359116952 residues

Total number of hits satisfying chosen parameters: 1595201

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

Published Applications AA:\*

1:	/cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubppaa/US09C_NEW_PUB.pep.*
12:	/cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubppaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1582	100.0	285 14 US-10-234-000-5	Sequence 5, Appl1
2	1582	100.0	285 14 US-10-411-120-60	Sequence 60, Appl1
3	1582	100.0	285 14 US-10-411-120-88	Sequence 88, Appl1
4	1582	100.0	285 15 US-10-621-787-2	Sequence 2, Appl1
5	1582	100.0	314 15 US-10-264-049-2619	Sequence 2619, Ap
6	1582	100.0	330 14 US-10-236-055A-6	Sequence 6, Appl1
7	1582	100.0	330 15 US-10-162-335-30	Sequence 30, Appl1
8	1578	99.7	331 15 US-10-220-120-354	Sequence 354, App
9	1492	94.3	294 14 US-10-236-055A-8	Sequence 8, Appl1
10	1487	94.0	294 10 US-09-866-050A-294	Sequence 294, App
11	1484	93.8	285 10 US-09-866-050A-382	Sequence 382, App
12	1154	72.9	217 9 US-09-770-906-2	Sequence 2, Appl1
13	981	62.0	198 10 US-09-866-050A-138	Sequence 138, App

14	898.5	56.8	289 14 US-10-234-000-15	Sequence 15, Appl1
15	898	56.8	289 10 US-09-866-050A-630	Sequence 630, App
16	895	56.6	273 17 US-10-758-846-60	Sequence 60, Appl1
17	895	56.6	289 14 US-10-411-120-59	Sequence 59, Appl1
18	895	56.6	289 15 US-10-451-168-106	Sequence 106, App
19	895	56.6	289 16 US-10-408-765A-1762	Sequence 1762, App
20	895	56.6	289 17 US-10-758-846-59	Sequence 59, Appl1
21	895	56.6	289 17 US-10-758-846-75	Sequence 75, Appl1
22	895	56.6	303 14 US-10-234-000-2	Sequence 2, Appl1
23	895	56.6	303 14 US-10-203-708-31	Sequence 31, Appl1
24	879.5	55.6	288 14 US-10-203-708-30	Sequence 30, Appl1
25	675	42.7	126 14 US-10-411-120-107	Sequence 107, App
26	671	42.4	125 17 US-10-758-846-101	Sequence 101, App
27	561.5	35.5	120 15 US-10-621-787-12	Sequence 12, Appl1
28	521.5	33.0	333 17 US-10-758-846-70	Sequence 70, Appl1
29	521.5	33.0	333 14 US-10-411-120-69	Sequence 69, Appl1
30	521.5	33.0	333 14 US-10-411-120-97	Sequence 97, Appl1
31	521.5	33.0	333 15 US-10-423-582-2	Sequence 2, Appl1
32	521.5	33.0	333 17 US-10-726-699-72	Sequence 72, Appl1
33	521.5	33.0	333 17 US-10-726-699-95	Sequence 95, Appl1
34	519.5	32.8	298 15 US-10-423-584-2	Sequence 2, Appl1
35	519.5	32.8	333 15 US-10-112-944-412	Sequence 412, App
36	519.5	32.8	333 17 US-10-758-846-10	Sequence 10, Appl1
37	519.5	32.8	680 14 US-10-177-293-59	Sequence 59, Appl1
38	519.5	32.8	680 15 US-10-058-270A-80	Sequence 80, Appl1
39	517.5	32.7	124 17 US-10-758-846-107	Sequence 107, App
40	516.5	32.6	680 14 US-10-295-027-110	Sequence 110, App
41	516.5	32.6	680 17 US-10-643-795A-118	Sequence 118, App
42	514.5	32.5	162 10 US-09-983-802-436	Sequence 436, App
43	513.5	32.5	162 10 US-09-984-490-436	Sequence 436, App
44	513.5	32.5	162 11 US-09-973-278-565	Sequence 565, App
45	513.5	32.5	162 14 US-10-411-120-101	Sequence 101, App
46	513.5	32.5	288 15 US-10-112-944-410	Sequence 410, App
47	507	32.0	288 17 US-10-758-846-4	Sequence 5, Appl1
48	506	32.0	269 17 US-10-758-846-5	Sequence 5, Appl1
49	506	32.0	508 17 US-10-758-846-90	Sequence 90, Appl1
50	482	30.5	98 9 US-09-770-906-4	Sequence 4, Appl1
51	479	30.3	482 17 US-10-758-846-92	Sequence 92, Appl1
52	466	29.5	493 17 US-10-758-846-89	Sequence 89, Appl1
53	460	29.1	459 16 US-10-408-765A-2554	Sequence 2554, Ap
54	457.5	28.9	459 17 US-10-758-846-28	Sequence 28, Appl1
55	457.5	28.9	477 17 US-10-758-846-27	Sequence 27, Appl1
56	457.5	28.9	477 17 US-10-758-846-91	Sequence 91, Appl1
57	456	28.8	476 17 US-10-758-846-93	Sequence 93, Appl1
58	453.5	28.7	481 17 US-10-758-846-88	Sequence 88, Appl1
59	451.5	28.5	438 17 US-10-758-846-78	Sequence 78, Appl1
60	450	28.4	438 15 US-10-138-588-42	Sequence 42, Appl1
61	446	28.2	703 14 US-10-219-449-4	Sequence 4, Appl1
62	444	28.1	717 14 US-10-219-449-2	Sequence 2, Appl1
63	444	28.1	733 15 US-10-138-588-40	Sequence 40, Appl1
64	444	28.1	638 13 US-10-001-887-108	Sequence 108, App
65	443	28.0	744 14 US-10-236-055A-32	Sequence 32, Appl1
66	438.5	27.7	703 14 US-10-411-120-65	Sequence 65, Appl1
67	436	27.6	744 14 US-10-301-822-39	Sequence 39, Appl1
68	436	27.6	744 17 US-10-758-846-71	Sequence 71, Appl1
69	436	27.6	717 17 US-10-758-846-25	Sequence 25, Appl1
70	434	27.4	744 14 US-10-236-055A-30	Sequence 30, Appl1
71	434	27.4	744 17 US-10-758-846-24	Sequence 24, Appl1
72	434	27.4	244 9 US-09-776-976-6	Sequence 6, Appl1
73	431.5	27.3	244 9 US-09-758-055-6	Sequence 6, Appl1
74	431.5	27.3	244 9 US-09-909-547-6	Sequence 6, Appl1
75	431.5	27.3	244 9 US-09-911-176B-48	Sequence 48, Appl1
76	431.5	27.3	244 14 US-10-180-762-51	Sequence 51, Appl1
77	431.5	27.3	244 14 US-10-331-814-6	Sequence 6, Appl1
78	431.5	27.3	244 14 US-10-234-000-4	Sequence 4, Appl1
79	431.5	27.3	244 14 US-10-360-186-51	Sequence 51, Appl1
80	431.5	27.3	244 14 US-10-321-164-3	Sequence 3, Appl1
81	431.5	27.3	244 14 US-10-189-493-1	Sequence 1, Appl1
82	431.5	27.3	244 14 US-10-197-293-3	Sequence 3, Appl1
83	431.5	27.3	244 14 US-10-325-717-1	Sequence 7, Appl1
84	431.5	27.3	244 14 US-10-411-120-67	Sequence 67, Appl1
85	431.5	27.3	244 15 US-10-379-747-6	Sequence 6, Appl1
86	431.5	27.3		

```
87 431.5 27.3 244 15 US-10-379-747-8 Sequence 8, Appli
88 431.5 27.3 244 15 US-10-379-747-10 Sequence 10, Appl
89 431.5 27.3 244 15 US-10-621-787-3 Sequence 10, Appl
90 431.5 27.3 244 15 US-10-285-833-6 Sequence 6, Appli
91 431.5 27.3 244 17 US-10-723-860-1048 Sequence 1048, Ap
92 431.5 27.3 250 15 US-10-379-747-12 Sequence 12, Appl
93 431 27.2 744 14 US-10-171-311-40 Sequence 40, Appl
94 431 27.2 744 14 US-10-301-822-37 Sequence 37, Appl
95 425 26.9 247 9 US-09-776-976-4 Sequence 4, Appli
96 425 26.9 247 9 US-09-758-055-4 Sequence 4, Appli
97 425 26.9 247 9 US-09-909-547-4 Sequence 4, Appli
98 425 26.9 247 14 US-10-231-814-4 Sequence 4, Appli
99 425 26.9 247 15 US-10-285-833-4 Sequence 4, Appli
100 424.5 26.8 244 14 US-10-376-460-6 Sequence 6, Appli
```

## ALIGNMENTS

```
RESULT 1
US-10-234-000-5
; Sequence 5, Application US/10234000
; Publication No. US20030129698A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN HOMOLOG ZACRP7
; FILE REFERENCE: 99-33
; CURRENT APPLICATION NUMBER: US/10/234,000
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US/09/577,298
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/136,289
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/145,589
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/158,448
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-234-000-5
```

```
Query Match 100.0%; Score 1582; DB 14; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.4e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPWVLLACALPCADADPLGAFARDFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
DB 1 MIPWVLLACALPCADADPLGAFARDFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
QY 61 GPFKDGQDGDHDDRGDSGSEGGPGRGNKGPCKAKAIGRAGRGKGVNGTGGK 120
DB 61 GPFKDGQDGDHDDRGDSGSEGGPGRGNKGPCKAKAIGRAGRGKGVNGTGGK 120
QY 121 GTGKKGPKKKGKGPGLPGPCSCSGHTKSAFSAVAVTKSYPRERLPKFKDKILNNEGSHY 180
DB 121 GTGKKGPKKKGKGPGLPGPCSCSGHTKSAFSAVAVTKSYPRERLPKFKDKILNNEGSHY 180
QY 181 NASSGKFCVCGVPGLIYFTYDITLANKHLAIGLVNNGQYRIRTPDANTGNHDVAGSTTLA 240
DB 181 NASSGKFCVCGVPGLIYFTYDITLANKHLAIGLVNNGQYRIRTPDANTGNHDVAGSTTLA 240
QY 241 LKQGDVWMLQIFYSEQNGLFYDPYWTDSLFTGFLIYADODDPNEV 285
DB 241 LKQGDVWMLQIFYSEQNGLFYDPYWTDSLFTGFLIYADODDPNEV 285
```

```
RESULT 2
US-10-411-120-60
```

```
; Sequence 60, Application US/10411120
; Publication No. US20030224425A1
; GENERAL INFORMATION:
; APPLICANT: Blondel et al.
; TITLE OF INVENTION: ACRP-30-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P7316P1
; CURRENT APPLICATION NUMBER: US/10/411,120
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/328,419
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: PCT/US02/32432
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 60
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-120-60
```

```
Query Match 100.0%; Score 1582; DB 14; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.4e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPWVLLACALPCADADPLGAFARDFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
DB 1 MIPWVLLACALPCADADPLGAFARDFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
QY 61 GPFKDGQDGDHDDRGDSGSEGGPGRGNKGPCKAKAIGRAGRGKGVNGTGGK 120
DB 61 GPFKDGQDGDHDDRGDSGSEGGPGRGNKGPCKAKAIGRAGRGKGVNGTGGK 120
QY 121 GTGKKGPKKKGKGPGLPGPCSCSGHTKSAFSAVAVTKSYPRERLPKFKDKILNNEGSHY 180
DB 121 GTGKKGPKKKGKGPGLPGPCSCSGHTKSAFSAVAVTKSYPRERLPKFKDKILNNEGSHY 180
QY 181 NASSGKFCVCGVPGLIYFTYDITLANKHLAIGLVNNGQYRIRTPDANTGNHDVAGSTTLA 240
DB 181 NASSGKFCVCGVPGLIYFTYDITLANKHLAIGLVNNGQYRIRTPDANTGNHDVAGSTTLA 240
QY 241 LKQGDVWMLQIFYSEQNGLFYDPYWTDSLFTGFLIYADODDPNEV 285
DB 241 LKQGDVWMLQIFYSEQNGLFYDPYWTDSLFTGFLIYADODDPNEV 285
```

## RESULT 3

```
US-10-411-120-88
; Sequence 88, Application US/10411120
; Publication No. US20030224425A1
; GENERAL INFORMATION:
; APPLICANT: Blondel et al.
; TITLE OF INVENTION: ACRP-30-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P7316P1
; CURRENT APPLICATION NUMBER: US/10/411,120
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/328,419
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: PCT/US02/32432
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 88
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-120-88
```

```
Query Match 100.0%; Score 1582; DB 14; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.4e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPWVLLACALPCADADPLGAFARDFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
```



```
Db 1 MIPVLLAALPCAADPULGAFARDFRKSPQLVCSLPGFQGPFGPPGAPGPGMMGRM 60
Qy 61 GPPGKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPGVNGTPEKH 120
Db 61 GPPGKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPGVNGTPEKH 120
Qy 121 GTPGKPKPKKGGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Db 121 GTPGKPKPKKGGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Qy 181 NASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRITFPDANTGNHDAVAGSTTLA 240
Db 181 NASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRITFPDANTGNHDAVAGSTTLA 240
Qy 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285
Db 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285

RESULT 4
US-10-621-787-2
; Sequence 2, Application US/10621787
; Publication No. US20040024187A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; TITLE OR INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG
; FILE REFERENCE: 99-08BD1
; CURRENT APPLICATION NUMBER: US/10/621,787
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: US 09/552,204
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: US 60/130,207
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-787-2

Query Match 100.0%; Score 1582; DB 15; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.4e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPVLLAALPCAADPULGAFARDFRKSPQLVCSLPGFQGPFGPPGAPGPGMMGRM 60
Db 1 MIPVLLAALPCAADPULGAFARDFRKSPQLVCSLPGFQGPFGPPGAPGPGMMGRM 60
Qy 61 GPPGKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPGVNGTPEKH 120
Db 61 GPPGKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPGVNGTPEKH 120
Qy 121 GTPGKPKPKKGGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Db 121 GTPGKPKPKKGGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Qy 181 NASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRITFPDANTGNHDAVAGSTTLA 240
Db 181 NASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRITFPDANTGNHDAVAGSTTLA 240
Qy 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285
Db 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285

RESULT 5
US-10-264-049-2619
; Sequence 2619, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
```

```
APPLICANT: Birse et al.
; TITLE OR INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2619
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2619

Query Match 100.0%; Score 1582; DB 15; Length 314;
Best Local Similarity 100.0%; Pred. No. 1.6e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPVLLAALPCAADPULGAFARDFRKSPQLVCSLPGFQGPFGPPGAPGPGMMGRM 60
Db 30 MIPVLLAALPCAADPULGAFARDFRKSPQLVCSLPGFQGPFGPPGAPGPGMMGRM 89
Qy 61 GPPGKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPGVNGTPEKH 120
Db 90 GPPGKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPGVNGTPEKH 149
Qy 121 GTPGKPKPKKGGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Db 150 GTPGKPKPKKGGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 209
Qy 181 NASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRITFPDANTGNHDAVAGSTTLA 240
Db 210 NASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRITFPDANTGNHDAVAGSTTLA 269
Qy 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285
Db 270 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 314

RESULT 6
US-10-236-055A-6
; Sequence 6, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Matson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Charles
; TITLE OR INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-055A-6

Query Match 100.0%; Score 1582; DB 14; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.7e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPVLLAALPCAADPULGAFARDFRKSPQLVCSLPGFQGPFGPPGAPGPGMMGRM 60
```

Db 46 MIPVLLACALPCAADPLGAFARDRFKSGPOLVCSLPGPQPGPGGAPGSGMMGRM 105  
Qy 61 GPRKDGQDGHDDRGSGSGEPGRIGNRGKPGPKGAKAIGRAGPRGPGVNGTGGK 120  
Db 106 GPRKDGQDGHDDRGSGSGEPGRIGNRGKPGPKGAKAIGRAGPRGPGVNGTGGK 165  
Qy 121 GTPKKGKPKGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPKFDKILNNEGSHY 180  
Db 166 GTPKKGKPKGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPKFDKILNNEGSHY 225  
Qy 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Db 226 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 285  
Qy 241 LKQDDEVWLOIFYSEONGLFYDPYWTDSLFTGFLIYADQDDPNEV 285  
Db 286 LKQDDEVWLOIFYSEONGLFYDPYWTDSLFTGFLIYADQDDPNEV 330

## RESULT 7

US-10-162-335-30  
; Sequence 30, Application US/10162335  
; Publication No. US2004009480A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David W.  
; APPLICANT: Baumgartner, Jason C.  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Gangoli, Baha A.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Gorman, Linda  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Hjal, Tord  
; APPLICANT: Kehuda, Rameah  
; APPLICANT: Li, Li  
; APPLICANT: Macdougall, John R.  
; APPLICANT: Malyanekar, Uriel M.  
; APPLICANT: Miller, Isabelle  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Paturajan, Meera  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Stone, David J.  
; APPLICANT: Szytek, Kimberly A.  
; APPLICANT: Vernet, Corine A. M.  
; APPLICANT: Voss, Edward Z.  
; APPLICANT: Zehnsen, Bryan D.  
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method  
; FILE REFERENCE: 21402-377 B  
; CURRENT APPLICATION NUMBER: US/10162,335  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: 60/295,607  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 60/295,661  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 60/296,404  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: 60/296,418  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: 60/297,414  
; PRIOR FILING DATE: 2001-06-11  
; PRIOR APPLICATION NUMBER: 60/297,567  
; PRIOR FILING DATE: 2001-06-12  
; PRIOR APPLICATION NUMBER: 60/298,285  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/298,556  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/299,949  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: 60/300,883  
; PRIOR FILING DATE: 2001-06-26

; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 201  
; SEQ ID NO 30  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-162-335-30

Query Match 100.0%; Score 1582; DB 15; Length 330;  
Best Local Similarity 100.0%; Pred. No. 1.7e-112;  
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPVLLACALPCAADPLGAFARDRFKSGPOLVCSLPGPQPGPGGAPGSGMMGRM 60  
Db 46 MIPVLLACALPCAADPLGAFARDRFKSGPOLVCSLPGPQPGPGGAPGSGMMGRM 105  
Qy 61 GPRKDGQDGHDDRGSGSGEPGRIGNRGKPGPKGAKAIGRAGPRGPGVNGTGGK 120  
Db 106 GPRKDGQDGHDDRGSGSGEPGRIGNRGKPGPKGAKAIGRAGPRGPGVNGTGGK 165  
Qy 121 GTPKKGKPKGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPKFDKILNNEGSHY 180  
Db 166 GTPKKGKPKGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPKFDKILNNEGSHY 225  
Qy 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Db 226 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 285  
Qy 241 LKQDDEVWLOIFYSEONGLFYDPYWTDSLFTGFLIYADQDDPNEV 285  
Db 286 LKQDDEVWLOIFYSEONGLFYDPYWTDSLFTGFLIYADQDDPNEV 330

## RESULT 8

US-10-220-120-354  
; Sequence 354, Application US/10220120  
; Publication No. US20040048253A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: PANZER, Scott R.  
; APPLICANT: SPIRO, Peter A.  
; APPLICANT: BANVILLE, Steven C.  
; APPLICANT: SHAH, Purvi  
; APPLICANT: CHALUP, Michael S.  
; APPLICANT: CHANG, Simon C.  
; APPLICANT: CHEN, Alice  
; APPLICANT: D'Sa, Steven A.  
; APPLICANT: AMSEY, Stefan  
; APPLICANT: DAHL, Christopher R.  
; APPLICANT: DAM, Tam C.  
; APPLICANT: DANIELS, Susan B.  
; APPLICANT: DUFOUR, Gerard B.  
; APPLICANT: FLORES, Vincent  
; APPLICANT: FONG, Willy T.  
; APPLICANT: GREENAWALT, Lila B.  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: JONES, Anissa L.  
; APPLICANT: LIT, Tommy F.  
; APPLICANT: ROSEBERRY, Ann M.  
; APPLICANT: ROSEN, Bruce H.  
; APPLICANT: RUSSO, Frank D.  
; APPLICANT: STOCKDREHER, Theresa K.  
; APPLICANT: DAFO, Abel  
; APPLICANT: WRIGHT, Rachel J.  
; APPLICANT: YAP, Pierre B.  
; APPLICANT: YU, Jimmy Y.  
; APPLICANT: BRADLEY, Diana L.  
; APPLICANT: BRATCHER, Shawn R.  
; APPLICANT: CHEN, Wensheng  
; APPLICANT: COHEN, Howard J.  
; APPLICANT: HODGSON, David M.  
; APPLICANT: LINCOLN, Stephen E.  
; APPLICANT: JACKSON, Stuart

TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: PT-1113 PCT  
 CURRENT APPLICATION NUMBER: US/10/220,120  
 CURRENT FILING DATE: 2002-08-26  
 PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;  
 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;  
 60/184,769; 60/184,768; 60/184,837; 60/184,841;  
 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;  
 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;  
 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;  
 60/205,324; 60/205,286  
 PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
 2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;  
 2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;  
 2000-05-17; 2000-05-17  
 NUMBER OF SEQ ID NOS: 422  
 SOFTWARE: PERL Program  
 SEQ ID NO 354  
 LENGTH: 331  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:414307.1.orf2:2000FEB01  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: 191  
 OTHER INFORMATION: unknown or other  
 US-10-220-120-354

Query Match 99.7%; Score 1578; DB 15; Length 331;  
 Best Local Similarity 99.6%; Pred. No. 3.3e-112;  
 Matches 284; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCADPLTGAFAARRDRKSPOLVCSIPGQGGPPGAGPSPGMMGRM 60  
 DB 47 MIPVLLACALPCADPLTGAFAARRDRKSPOLVCSIPGQGGPPGAGPSPGMMGRM 106  
 QY 61 GPGKDGDDGHDGRGSGEGRPPGRTGNRKGPPKGAAGIAGRAPPGKGVNGTPEGK 120  
 DB 107 GPGKDGDDGHDGRGSGEGRPPGRTGNRKGPPKGAAGIAGRAPPGKGVNGTPEGK 166  
 QY 121 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNMEGHY 180  
 DB 167 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNMEGHY 226  
 QY 181 NASSGKRVCGVPGIYFTYDITLANKKLAIGLVHNGQRIITPDANTGNHVAAGSTTLA 240  
 DB 227 NASSGKRVCGVPGIYFTYDITLANKKLAIGLVHNGQRIITPDANTGNHVAAGSTTLA 286  
 QY 241 LKQGDVWLQIFYSQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285  
 DB 287 LKQGDVWLQIFYSQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 331

RESULT 9  
 US-10-236-055A-8  
 Sequence 8, Application US/10236055A  
 Publication No. US20030134328A1  
 GENERAL INFORMATION:  
 APPLICANT: Basham, Beth E.  
 APPLICANT: Forsythe, Ian  
 APPLICANT: Gorman, Daniel M.  
 APPLICANT: Matteson, Jeanne  
 APPLICANT: Moshrefi, Chiriel  
 TITLE OF INVENTION: MAMMALIAN GENES, RELATED REAGENTS  
 FILE REFERENCE: DX01343K  
 CURRENT APPLICATION NUMBER: US/10/236,055A  
 CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/317,988  
 PRIOR FILING DATE: 2001-09-06  
 NUMBER OF SEQ ID NOS: 36  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 8  
 LENGTH: 294  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-10-236-055A-8

Query Match 94.3%; Score 1492; DB 14; Length 294;  
 Best Local Similarity 94.0%; Pred. No. 1e-105;  
 Matches 269; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCADPLTGAFAARRDRKSPOLVCSIPGQGGPPGAGPSPGMMGRM 60  
 DB 10 MISVLLACALPCADPLTGAFAARRDRKSPOLVCSIPGQGGPPGAGPSPGMMGRM 69  
 QY 61 GPGKDGDDGHDGRGSGEGRPPGRTGNRKGPPKGAAGIAGRAPPGKGVNGTPEGK 120  
 DB 70 GPGKDGDDGHDGRGSGEGRPPGRTGNRKGPPKGAAGIAGRAPPGKGVNGTPEGK 129  
 QY 121 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNMEGHY 180  
 DB 130 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNMEGHY 189  
 QY 181 NASSGKRVCGVPGIYFTYDITLANKKLAIGLVHNGQRIITPDANTGNHVAAGSTTLA 240  
 DB 190 NASSGKRVCGVPGIYFTYDITLANKKLAIGLVHNGQRIITPDANTGNHVAAGSTTLA 249  
 QY 241 LKQGDVWLQIFYSQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285  
 DB 250 LKQGDVWLQIFYSQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 294

RESULT 10  
 US-09-866-050A-294  
 Sequence 294, Application US/09866050A  
 Publication No. US20030040471A1  
 GENERAL INFORMATION:  
 APPLICANT: Watson, James D.  
 APPLICANT: Strachan, Jorna  
 APPLICANT: Sleeman, Matthew  
 APPLICANT: Ornstet, Rene  
 APPLICANT: Murlison, James G.  
 TITLE OF INVENTION: Compositions Isolated From Skin Cells  
 FILE REFERENCE: 11000.1011c4V  
 CURRENT APPLICATION NUMBER: US/09/866,050A  
 CURRENT FILING DATE: 2001-05-24  
 NUMBER OF SEQ ID NOS: 725  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 294  
 LENGTH: 294  
 TYPE: PRT  
 ORGANISM: Rat  
 US-09-866-050A-294

Query Match 94.0%; Score 1487; DB 10; Length 294;  
 Best Local Similarity 94.0%; Pred. No. 2.5e-105;  
 Matches 268; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCADPLTGAFAARRDRKSPOLVCSIPGQGGPPGAGPSPGMMGRM 60  
 DB 10 MISVLLACALPCADPLTGAFAARRDRKSPOLVCSIPGQGGPPGAGPSPGMMGRM 69  
 QY 61 GPGKDGDDGHDGRGSGEGRPPGRTGNRKGPPKGAAGIAGRAPPGKGVNGTPEGK 120  
 DB 70 GPGKDGDDGHDGRGSGEGRPPGRTGNRKGPPKGAAGIAGRAPPGKGVNGTPEGK 129  
 QY 121 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNMEGHY 180

Db 130 GIPGKGPCKGKGPGLPGPCSCSSRAKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 189  
Qy 181 NASGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Db 190 NASGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 249  
Qy 241 LKQDEVMWLQIFYSEONGLFYDPYMTDSLFTGFLIYADODDPNEV 285  
Db 250 LKQDEVMWLQIFYSEONGLFYDPYMTDSLFTGFLIYADODDPNEV 294

RESULT 11  
US-09-866-050A-382

; Sequence 382, Application US/09866050A  
; Publication No. US20030040471A1  
; GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Muriagon, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c4U  
; CURRENT APPLICATION NUMBER: US/09/866,050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 382  
; LENGTH: 285  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-866-050A-382

Query Match 93.8%; Score 1484; DB 10; Length 285;  
Best Local Similarity 93.7%; Pred. No. 4,1e-105;  
Matches 267; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MIPWVLLACALPCADBLGAFARDFRKSPOLVCSLPGPQGPBPAGAPSGMMGRM 60  
Db 1 MISWMLLACALPCADBLGAFARDFRKSPOLVCSLPGPQGPBPAGAPSGMMGRM 60  
Qy 61 GPFKQGDQDHDGRGSGSEGGPPGRTGNKGPCKKAGALGAGRGPAGVGTGKH 120  
Db 61 GPFKQGDQDHDGRGSGSEGGPPGRTGNKGPCKKAGALGAGRGPAGVGTGKH 120  
Qy 121 GTPGKGPCKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 180  
Db 121 GTPGKGPCKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 180  
Qy 181 NASGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Db 181 NASGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Qy 241 LKQDEVMWLQIFYSEONGLFYDPYMTDSLFTGFLIYADODDPNEV 285  
Db 241 LKQDEVMWLQIFYSEONGLFYDPYMTDSLFTGFLIYADODDPNEV 285

RESULT 12  
US-09-770-906-2

; Sequence 2, Application US/09770906  
; Patent No. US2001000905A1  
; GENERAL INFORMATION:

; APPLICANT: Hensley, Preston  
; APPLICANT: HU, ERDING  
; APPLICANT: SMITH, RANDALL, FORREST  
; APPLICANT: ZHU, YUAN  
; TITLE OF INVENTION: A HOMOLOG OF ACRP30 (30 KD  
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT-RELATED PROTEIN)  
; FILE REFERENCE: GP-70435-C1  
; CURRENT APPLICATION NUMBER: US/09/770,906

; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/086,562  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: 09/162,352  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 217  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-770-906-2

Query Match 72.9%; Score 1154; DB 9; Length 217;  
Best Local Similarity 76.1%; Pred. No. 4,3e-80;  
Matches 217; Conservative 0; Mismatches 0; Indels 68; Gaps 1;

Qy 1 MIPWVLLACALPCADBLGAFARDFRKSPOLVCSLPGPQGPBPAGAPSGMMGRM 60  
Db 1 MISWMLLACALPCADBLGAFARDFRKSPOLVCSLPGPQGPBPAGAPSGMMGRM 60  
Qy 61 GPFKQGDQDHDGRGSGSEGGPPGRTGNKGPCKKAGALGAGRGPAGVGTGKH 120  
Db 61 GPFKQGDQDHDGRGSGSEGGPPGRTGNKGPCKKAGALGAGRGPAGVGTGKH 120  
Qy 121 GTPGKGPCKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 180  
Db 121 GTPGKGPCKGKGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 180  
Qy 181 NASGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Db 181 NASGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Qy 241 LKQDEVMWLQIFYSEONGLFYDPYMTDSLFTGFLIYADODDPNEV 285  
Db 241 LKQDEVMWLQIFYSEONGLFYDPYMTDSLFTGFLIYADODDPNEV 285

RESULT 13  
US-09-866-050A-138

; Sequence 138, Application US/09866050A  
; Publication No. US20030040471A1  
; GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Muriagon, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c4U  
; CURRENT APPLICATION NUMBER: US/09/866,050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 138  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-866-050A-138

Query Match 62.0%; Score 981; DB 10; Length 198;  
Best Local Similarity 92.6%; Pred. No. 5,9e-67;  
Matches 175; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MIPWVLLACALPCADBLGAFARDFRKSPOLVCSLPGPQGPBPAGAPSGMMGRM 60  
Db 1 MISWMLLACALPCADBLGAFARDFRKSPOLVCSLPGPQGPBPAGAPSGMMGRM 60  
Qy 61 GPFKQGDQDHDGRGSGSEGGPPGRTGNKGPCKKAGALGAGRGPAGVGTGKH 120  
Db 61 GPFKQGDQDHDGRGSGSEGGPPGRTGNKGPCKKAGALGAGRGPAGVGTGKH 120

Qy 121 GTPGKGPCKGKGGPGLPGPCSGSGHTKSAFSAVATKSYPRERLPFKDKILNNEGHTY 180  
 Db 130 GTPGKGPCKGKGGPGLPGPCSSSRKSAFSAVATKSYPRERLPFKDKILNNEGHTY 189  
 Qy 181 NASSGKFVC 189  
 Db 190 NASSGKFVC 198

## RESULT 14

US-10-234-000-15  
 ; Sequence 15, Application US/10234000  
 ; Publication No. US20030129698A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Piddington, Christopher S.  
 ; APPLICANT: Sheppard, Paul O.  
 ; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN HOMOLOG ZACRP7  
 ; FILE REFERENCE: 99-31  
 ; CURRENT APPLICATION NUMBER: US/10/234,000  
 ; CURRENT FILING DATE: 2002-08-30  
 ; PRIOR APPLICATION NUMBER: US/09/577,298  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/136,289  
 ; PRIOR FILING DATE: 1999-05-27  
 ; PRIOR APPLICATION NUMBER: 60/145,589  
 ; PRIOR FILING DATE: 1999-07-22  
 ; PRIOR APPLICATION NUMBER: 60/158,448  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 15  
 ; LENGTH: 289  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-10-234-000-15

Query Match 56.8%; Score 898.5; DB 14; Length 289;  
 Best Local Similarity 62.7%; Pred. No. 1.7e-60;

Matches 163; Conservative 32; Mismatches 62; Indels 3; Gaps 1;

Qy 24 RRPFRKG---SPQLVSLPFPQGPFGAPGSPGMMGRMGFPKGQDGDHGDGRDSCGE 80  
 Db 19 RANQAKESYSRVCISIPGLPGPPGANGSPGPHGRIGLPGRDGRDGRKGEKRGKT 78  
 Qy 81 EGPPRTGNRGKPPKKAAGIAGAPRGPGVNGTPGKHGTPEKKGPKGKGGEPGLPGP 140  
 Db 79 AGLKGTGPIGLAGEKGDQGTGKKGPIGPEGEKGVGPAGPPGPKGDRDQDGPGLPGV 138  
 Qy 141 CSGSGHTKSAFSAVATKSYPRERLPFKDKILNNEGHTYASGKTVCGVPGIYFTYD 200  
 Db 139 CRGGSIVLKSAFSAVATKSYPRERLPFKDKILNNEGHTYASGKTVCAFPGLIYFSYD 198  
 Qy 201 ITLANKHLAIGLVNAGQYRIRTPDANTGNHVDVAGSSTIALKOGDEVWLOIFYSQNGLF 260  
 Db 199 ITLANKHLAIGLVNAGQYRIRTPDANTGNHVDVAGSSTIVYIQLPDEVWLEIFPDQNGLF 258  
 Qy 261 YDPYMTSLFTGFLIYADQD 280  
 Db 259 SDPGMADSLFSGFLIYVTD 278

## RESULT 15

US-09-866-050A-630  
 ; Sequence 630, Application US/09866050A  
 ; Publication No. US20030040471A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Watson, James D.  
 ; APPLICANT: Strachan, Lorna  
 ; APPLICANT: Sleeman, Matthew  
 ; APPLICANT: Onrust, Rene  
 ; APPLICANT: Murison, James G.  
 ; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
 ; TITLE OF INVENTION: and Methods for Their Use  
 ; FILE REFERENCE: 11000.1011c4U  
 ; CURRENT APPLICATION NUMBER: US/09/866,050A  
 ; CURRENT FILING DATE: 2001-05-24  
 ; NUMBER OF SEQ ID NOS: 725  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 630  
 ; LENGTH: 289  
 ; TYPE: PRT  
 ; ORGANISM: Rat  
 ; US-09-866-050A-630

Query Match 56.8%; Score 898; DB 10; Length 289;  
 Best Local Similarity 64.0%; Pred. No. 1.9e-60;

Matches 160; Conservative 31; Mismatches 59; Indels 0; Gaps 0;

Qy 31 SPQLVSLPFPQGPFGAPGSPGMMGRMGFPKGQDGDHGDGRDSCGERGPGRTGNR 90  
 Db 29 SPRVCISIPGLPGPPGANGSPGPHGRIGLPGRDGRDGRKGEKGTAGLKGKGPL 88  
 Qy 91 GKPPKKAAGIAGAPRGPGVNGTPGKHGTPEKKGPKGKGGEPGLPGPCSGSGHTKS 150  
 Db 89 GLAGEKGDQGTGKKGPIGPEGEKGVGPAGPPGPKGDRDQDGPGLPGVCRGCSIVLKS 148  
 Qy 151 AFSVAATKSYPRERLPFKDKILNNEGHTYASGKTVCGVPGIYFTYDITLANKHLAI 210  
 Db 149 AFSVGIITTSYPRERLPFKDKILNNEGHTYASGKTVCAFPGLIYFSYDITLANKHLAI 208  
 Qy 211 GLVNGQYRIRTPDANTGNHVDVAGSSTIALKOGDEVWLOIFYSQNGLFYDPYMTSLF 270  
 Db 209 GLVNGQYRIRTPDANTGNHVDVAGSSTIVYIQLPDEVWLEIFPDQNGLSDPGMADSLF 268  
 Qy 271 TGFLLIYADQD 280  
 Db 269 SGFLIYVTD 278

Search completed: December 25, 2004, 09:22:17  
 Job time : 505.456 secs

This Page Blank (USD)  
Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 25, 2004, 06:40:18 ; Search time 33.6487 Seconds  
(without alignments)  
330.491 Million cell updates/sec

Title: US-10-621-787-5  
Perfect score: 57  
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXGXXYFX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 100 summaries

## Database :

1: A\_Geneseq\_23Sep04:\*  
2: geneseqp1980s:\*  
3: geneseqp1990s:\*  
4: geneseqp2000s:\*  
5: geneseqp2001s:\*  
6: geneseqp2002s:\*  
7: geneseqp2003as:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	22	38.6	119 8 ADK66184	Adk66184 Mouse phx
2	22	38.6	136 8 ADK66189	Adk66189 Human phx
3	22	38.6	163 2 AAW70458	Aaw70458 Human T1-
4	22	38.6	188 4 AAM93624	Aam93624 Human pol
5	22	38.6	188 8 ADL31427	Adl31427 Human pro
6	22	38.6	215 2 AAW70459	Aaw70459 Human T1-
7	22	38.6	215 5 ABP65162	Abp65162 Hypoxia-r
8	22	38.6	219 4 ABB11780	Abb11780 Human T1
9	22	38.6	224 4 AAM93844	Aam93844 Human pol
10	22	38.6	224 4 AAB92820	Aab92820 Human pro
11	22	38.6	224 4 ADL31890	Adl31890 Human pro
12	22	38.6	231 6 ABR41674	Ab41674 Human DIT
13	22	38.6	266 4 AAG82486	Aag82486 S. epider
14	22	38.6	268 7 ABO69527	Ab69527 Pseudomon
15	22	38.6	292 5 ABP39297	Abp39297 Staphyloc
16	22	38.6	296 6 ABU39079	Abu39079 Protein e
17	22	38.6	302 6 ABU49473	Abu49473 Protein e
18	22	38.6	305 6 ADA34782	Ada34782 Actinobac
19	22	38.6	404 6 ADA36895	Ada36895 Actinobac
20	22	38.6	485 6 ABM67542	Abm67542 Phototrab
21	22	38.6	530 8 ADP04497	Adp04497 Sea equit
22	22	38.6	540 4 AAU57054	Aau57054 Propionib
23	22	38.6	540 6 ABM53573	Abm53573 Proteinib
24	22	38.6	3705 6 ABU50020	Abu50020 Protein e
25	21	36.8	20 4 AAB73384	Aab73384 T. harzia

26	21	36.8	53 8 ABO55435	Ab055435 Human gen
27	21	36.8	111 2 AAW89961	Aaw89961 Antigen f
28	21	36.8	128 2 ABG76563	Abg76563 HCV El an
29	21	36.8	151 2 AAY37071	Aay37071 Protein w
30	21	36.8	154 6 ABP75824	Abp75824 Human sec
31	21	36.8	156 2 AAY35280	Aay35280 Chlamydia
32	21	36.8	156 6 ABM26784	Abm26784 Protein e
33	21	36.8	161 6 ABM69553	Abm69553 Phototrab
34	21	36.8	169 5 ABP27601	Abp27601 Streptoco
35	21	36.8	183 5 ABP69421	Abp69421 Human pol
36	21	36.8	201 4 AAB74750	Aab74750 Human sec
37	21	36.8	201 4 AAB74759	Aab74759 Human sec
38	21	36.8	201 4 AAB74760	Aab74760 Human sec
39	21	36.8	201 5 ABG65326	Abg65326 Human alb
40	21	36.8	201 5 ABG65324	Abg65324 Human alb
41	21	36.8	201 7 ADD44752	Add44752 Human pro
42	21	36.8	201 7 ADD44746	Add44746 Rat Prote
43	21	36.8	201 7 ADD44750	Add44750 Rat Prote
44	21	36.8	201 7 ADD44748	Add44748 Human pro
45	21	36.8	201 7 ADJ68911	Adj68911 Human hea
46	21	36.8	201 7 ADP65317	Adp65317 Human coa
47	21	36.8	201 8 ADL78593	Adl78593 Albumin f
48	21	36.8	201 8 ADL78591	Adl78591 Albumin f
49	21	36.8	201 8 ADL78594	Adl78594 Albumin f
50	21	36.8	201 8 ADN04050	Adn04050 Antipepti
51	21	36.8	208 5 ABP69422	Abp69422 Human pol
52	21	36.8	210 2 AAW20426	Aaw20426 H. pylori
53	21	36.8	215 2 AAW20968	Aaw20968 H. pylori
54	21	36.8	238 6 ABU30638	Abu30638 Protein e
55	21	36.8	252 7 ABO82067	Ab082067 Pseudomon
56	21	36.8	254 4 AAU31151	Aau31151 Novel hum
57	21	36.8	254 7 ADB10062	Adb10062 Novel pro
58	21	36.8	271 3 AAG51318	Aag51318 Arabidops
59	21	36.8	271 3 AAG07614	Aag07614 Arabidops
60	21	36.8	277 3 AAG51317	Aag51317 Arabidops
61	21	36.8	277 3 AAG07613	Aag07613 Arabidops
62	21	36.8	277 3 ADD30338	Add30338 Plant yie
63	21	36.8	277 8 ADI44261	Adi44261 Plant tra
64	21	36.8	280 5 ABP29735	Abp29735 Streptoco
65	21	36.8	281 4 ABG23769	Abg23769 Novel hum
66	21	36.8	295 3 AAG51316	Aag51316 Arabidops
67	21	36.8	296 3 AAG07612	Aag07612 Arabidops
68	21	36.8	296 3 AAY91317	Aay91317 Group B S
69	21	36.8	300 5 ABP26518	Abp26518 Streptoco
70	21	36.8	307 3 AAB18332	Aab18332 Plasmodu
71	21	36.8	336 5 ABP62884	Abp62884 Human pol
72	21	36.8	348 4 ADM25554	Adm25554 Hyperther
73	21	36.8	382 4 ABB60990	Abb60990 Drosophi
74	21	36.8	382 4 AAB79473	Aab79473 Corynebac
75	21	36.8	408 4 AAG92180	Aag92180 C glutami
76	21	36.8	441 7 ABO63535	Ab063535 Klebsiell
77	21	36.8	450 7 AAU72909	Aau72909 Human met
78	21	36.8	473 5 ABB98126	Abb98126 Human PPM
79	21	36.8	473 3 AAY77945	Aay77945 A. thalita
80	21	36.8	508 3 AAG45662	Aag45662 Arabidops
81	21	36.8	729 3 AAG45661	Aag45661 Arabidops
82	21	36.8	783 3 AAG45660	Aag45660 Arabidops
83	21	36.8	803 4 ABB64390	Abb64390 Drosophi
84	21	36.8	1032 7 ABB62164	Abb62164 Drosophi
85	21	36.8	1272 7 ADH87086	Adh87086 Enterococ
86	21	36.8	1416 5 AAY00211	Aay00211 Enterococ
87	21	36.8	1416 5 ABP43430	Abp43430 E faecali
88	21	36.8	1416 6 ABU88458	Abu88458 E. faecal
89	21	36.8	1416 6 ABU13709	Abu13709 Enterococ
90	21	36.8	1448 6 AAY00210	Aay00210 Enterococ
91	21	36.8	1448 6 ABB43429	Abb43429 E faecali
92	21	36.8	1448 6 ABU88457	Abu88457 E. faecal
93	21	36.8	1448 6 ABU13708	Abu13708 Enterococ
94	21	36.8	1448 6 AAU21262	Aau21262 Human nov
95	20	35.1	53 8 ABO54754	Ab054754 Human gen
96	20	35.1	55 8 AAG22312	Aag22312 Zea may
97	20	35.1	94 3 AAG22311	Aag22311 Zea may
98	20	35.1		

99 20 35.1 118 4 AAm94043  
100 20 35.1 122 8 ADN47176

AAm94043 Human rep  
ADN47176 Thermococ

## ALIGNMENTS

## RESULT 1

ADK66184  
ID ADK66184 standard; protein; 119 AA.

AC ADK66184;

DT 06-MAY-2004 (first entry)

DE Mouse pfx-6 protein.

XX Novel gene; pain; neuropathic pain; shingles pain;

KM post-herpetic neuralgia; gene therapy; mouse.

OS Mus musculus.

XX Key Location/Qualifiers

FT Misc-difference 110 /note= "Encoded by TGC"

FT Misc-difference 117 /note= "Encoded by GTG"

XX US2004019006-A1.

XX 29-JAN-2004.

PF 09-MAY-2003; 2003US-00434156.

PR 10-MAY-2002; 2002US-0378955P.

XX (HAYA/) HAYASHIZAKI Y.

PA (KAMI/) KAMIYA M.

PA (SUZU/) SUZUKI T.

PA (HIRO/) HIROKAWA H.

PA (NAKA/) NAKAO K.

PA (TANA/) TANAKA T.

PI Hayashizaki Y, Kamiya M, Suzuki T, Hirokawa H, Nakao K, Tanaka T;

XX WPI; 2004-122083/12.

DR N-PSDB; ADK66173.

XX New isolated polynucleotide, useful for useful for treating or preventing  
PT pain, e.g. neuropathic pain, preferably shingles pain or post-herpetic  
PT neuralgia.

PS Example 9; SEQ ID NO 17; 53pp; English:

XX The present invention provides novel genes relating to pain. The  
CC invention is useful for treating and preventing pain such as neuropathic  
CC pain, shingles pain and post-herpetic neuralgia. The invention is also  
CC useful in gene therapy. The present sequence is mouse pfx-6 protein.

XX Sequence 119 AA;

SQ Query Match 38.6%; Score 22; DB 8; Length 119;

Best Local Similarity 36.4%; Pred. No. 2.8e+02; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYYF 29

DB 10 FTASRNGTYKF 20

## RESULT 2

ADK66189

ID ADK66189 standard; protein; 136 AA.

XX ADK66189;

AC 06-MAY-2004 (first entry)

DE Human pfx-6 protein.

XX Novel gene; pain; neuropathic pain; shingles pain;

KM post-herpetic neuralgia; gene therapy; human.

OS Homo sapiens.

XX US2004019006-A1.

XX 29-JAN-2004.

PF 09-MAY-2003; 2003US-00434156.

PR 10-MAY-2002; 2002US-0378955P.

XX (HAYA/) HAYASHIZAKI Y.

PA (KAMI/) KAMIYA M.

PA (SUZU/) SUZUKI T.

PA (HIRO/) HIROKAWA H.

PA (NAKA/) NAKAO K.

PA (TANA/) TANAKA T.

PI Hayashizaki Y, Kamiya M, Suzuki T, Hirokawa H, Nakao K, Tanaka T;

XX WPI; 2004-122083/12.

DR N-PSDB; ADK66178.

XX New isolated polynucleotide, useful for useful for treating or preventing  
PT pain, e.g. neuropathic pain, preferably shingles pain or post-herpetic  
PT neuralgia.

PS Claim 10; SEQ ID NO 22; 53pp; English.

XX The present invention provides novel genes relating to pain. The  
CC invention is useful for treating and preventing pain such as neuropathic  
CC pain, shingles pain and post-herpetic neuralgia. The invention is also  
CC useful in gene therapy. The present sequence is human pfx-6 protein.

XX Sequence 136 AA;

SQ Query Match 38.6%; Score 22; DB 8; Length 136;

Best Local Similarity 36.4%; Pred. No. 3.2e+02; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYYF 29

DB 10 FTASRNGTYKF 20

## RESULT 3

AAW70458

ID AAW70458 standard; protein; 163 AA.

XX AAW70458;

DT 10-DEC-1998 (first entry)

DE Human T1-receptor ligand III splice variant 1.

XX Human T1-receptor ligand III; T1-R ligand III; atherosclerosis;

KM autoimmune disease; inflammation; metabolic dysfunction;

KW immune-regulated disorder.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1.24 /note= "Signal peptide"



PT Protein 25.163  
 FT /note="Tl-R ligand III splice variant 1"  
 XX  
 PN W09838311-A1.  
 XX  
 PD 03-SEP-1998.  
 XX  
 PF 26-FEB-1998; 98WO-US003483.  
 XX  
 PR 28-FEB-1997; 97US-0039483P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI N1 J, Gentz RL, Ruben SM;  
 XX  
 DR WPI; 1998-495398/42.  
 XX  
 DR N-PSDB; AAV33461.  
 XX  
 PT New human Tl-receptor ligand III splice variant (s) - useful for, e.g.  
 PT detecting and treating immune system related disorders such as cancer and  
 PT inflammation.  
 XX  
 PS Claim 1; Fig 1; 115p; English.  
 XX  
 CC The invention provides novel human Tl-receptor ligand III (Tl-R ligand  
 CC III) splice variants and antibodies raised against these proteins. The  
 CC present sequence represents the Tl-R ligand III splice variant 1 protein  
 CC sequence. The Tl-R ligand III splice variants are claimed useful for  
 CC screening agonists and antagonists. They are also claimed useful for  
 CC treating disorders such as atherosclerosis, autoimmune disease,  
 CC inflammation, metabolic dysfunction and immune-regulated disorders  
 XX  
 SQ Sequence 163 AA;

Query Match 38.6%; Score 22; DB 2; Length 163;  
 Best Local Similarity 36.4%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXKX 29  
 DB 89 FTASKNGTYKF 99

RESULT 4  
 AAM93624  
 ID AAM93624 standard; protein; 188 AA.  
 XX  
 AC AAM93624;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polypeptide, SEQ ID NO: 3460.  
 XX  
 KW Human, full length cDNA; cDNA synthesis; oligo-capping.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP130094-A2.  
 XX  
 PD 05-SEP-2001.  
 XX  
 PF 07-JUL-2000; 2000EP-00114089.  
 XX  
 PR 08-JUL-1999; 99JP-00194486.  
 XX  
 PR 11-JAN-2000; 2000JP-00118774.  
 XX  
 PR 02-MAY-2000; 2000JP-00183765.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 DR WPI; 2001-524255/58.

DR N-PSDB; AAK94557.  
 XX  
 PT 830 Primers useful for synthesizing full length cDNA clones and their use  
 PT in genetic manipulation.  
 XX  
 PS Claim 8; SEQ ID NO 3460; 1380bp + Sequence Listing; English.  
 XX  
 CC The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been isolated  
 CC and nucleotide sequences of 5' - and 3' - ends of the cDNA molecules have  
 CC been determined. Primers for synthesizing the full length cDNA are useful  
 CC for clarifying the function of the protein encoded by the cDNA. The full  
 CC length clones were obtained by construction of full length enriched cDNA  
 CC libraries that were synthesised by the oligo-capping method. The primers  
 CC enable the production of the full length cDNA easily without any special  
 CC methods. The present sequence is a polypeptide encoded by a full length  
 CC human cDNA of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in CD-ROM  
 CC format directly from EPO  
 XX  
 SQ Sequence 188 AA;

Query Match 38.6%; Score 22; DB 4; Length 188;  
 Best Local Similarity 36.4%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXKX 29  
 DB 98 FTASKNGTYKF 108

RESULT 5  
 ADL31427  
 ID ADL31427 standard; protein; 188 AA.  
 XX  
 AC ADL31427;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Human protein encoded by a full length cDNA clone SegID 3460.  
 XX  
 KW human; medicine; signal transduction; glycoprotein; transcription;  
 XX oligo-capping method.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1396543-A2.  
 XX  
 PD 10-MAR-2004.  
 XX  
 PF 07-JUL-2000; 2003EP-00025638.  
 XX  
 PR 08-JUL-1999; 99JP-00194486.  
 XX  
 PR 11-JAN-2000; 2000JP-00118774.  
 XX  
 PR 02-MAY-2000; 2000JP-00183865.  
 XX  
 PR 07-JUL-2000; 2000EP-00114089.  
 XX  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 DR WPI; 2004-204755/20.  
 XX  
 DR N-PSDB; ADL31426.  
 XX  
 PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
 PT length human cDNAs.  
 XX  
 PS Example 1; SEQ ID NO 3460; 1340bp; English.  
 XX  
 CC This invention relates to a novel primers useful for synthesizing full  
 CC length cDNA molecules that encode human proteins. Specifically, it refers  
 CC to secretory or membrane proteins that are potential therapeutic agents/

CC target molecules in the field of medicine, and in particular genes  
 CC encoding proteins that are associated with signal transduction.  
 CC glycoproteins and transcription. The present invention describes a method  
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
 CC ends using the oligo-capping method. This polypeptide sequence is a full  
 CC length human protein of the invention.

XX Sequence 188 AA;

SQ Query Match 38.6%; Score 22; DB 8; Length 188;

Best Local Similarity 36.4%; Pred. No. 4.4e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYXF 29

Db 98 FTASKNGTYKF 108

RESULT 6

AAW70459 ID AAW70459 standard; protein; 215 AA.

XX AAW70459;

XX 10-DEC-1998 (first entry)

XX Human T1-receptor ligand III splice variant 2.

XX Human T1-receptor ligand III; T1-R ligand III; atherosclerosis;  
 KM autoimmune disease; inflammation; metabolic dysfunction;  
 KM immune-regulated disorder.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..24

XX Protein /note= "Signal peptide"

XX W09838311-A1.

XX 03-SEP-1998.

XX 26-FEB-1998; 98WO-US003483.

XX 28-FEB-1997; 97US-0039483P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Gentz RL, Ruben SM;

XX WPI; 1998-495398/42.

XX N-PSDB; AAV33462.

XX New human T1-receptor ligand III splice variant(s) - useful for, e.g.  
 PT detecting and treating immune system related disorders such as cancer and  
 PT inflammation.

XX Claim 1; Fig 2; 115pp; English.

XX The invention provides novel human T1-receptor ligand III (T1-R ligand  
 CC III) splice variants and antibodies raised against these proteins. The  
 CC present sequence represents the T1-R ligand III splice variant 2 protein  
 CC sequence. The T1-R ligand III splice variants are claimed useful for  
 CC screening agonists and antagonists. They are also claimed useful for  
 CC treating disorders such as atherosclerosis, autoimmune disease,  
 CC inflammation, metabolic dysfunction and immune-regulated disorders

XX Sequence 215 AA;

Query Match 38.6%; Score 22; DB 2; Length 215;  
 Best Local Similarity 36.4%; Pred. No. 5e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYXF 29

Db 89 FTASKNGTYKF 99

RESULT 7

ID ABP65162 standard; protein; 215 AA.

XX ABP65162;

XX 12-NOV-2002 (first entry)

XX Hypoxia-regulated protein #36.

XX Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;  
 KM antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;  
 KM hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;

KM ischaemic condition; reperfusion injury; retinopathy; neonatal stress;  
 KM preclampsia; atherosclerosis; inflammatory condition; wound healing;  
 KM inflammation; erythropoiesis; hair loss; human.

XX Homo sapiens.

XX W0200246465-A2.

XX 13-JUN-2002.

XX 10-DEC-2001; 2001WO-GB005458.

XX 08-DEC-2000; 2000GB-00030076.

XX 08-FEB-2001; 2001GB-00003156.

XX 25-OCT-2001; 2001GB-00025666.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX White J, Mundy CR, Ward NR, Krige D, Kingman SM, Harris RA;

XX Rayner KN;

XX WPI; 2002-627238/67.

XX Identifying a gene involved in disease for treating hypoxia-regulated  
 PT conditions, comprises comparing the transcriptome/proteome of two cell  
 PT types under different conditions and identifying a differentially  
 PT regulated gene.

XX Claim 35; Page 355; 538pp; English.

XX The present invention relates to methods for identifying genes and  
 CC proteins that are implicated in a specific disease or physiological  
 CC condition. The method comprises comparing the transcriptome/proteome of a  
 CC specialised cell type implicated in a disease or condition with that of a  
 CC second specialised cell type, under two experimental conditions, and  
 CC identifying a gene that is differentially regulated in the two  
 CC specialised cell types under experimental conditions. ABV7873-ABV78116  
 CC and ABP65061-ABP65257 were identified using the methods of the invention.  
 CC The coding sequences and proteins are useful for treating a disease in a  
 CC patient, for manufacture of a medicament for treating hypoxia-regulated  
 CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,  
 CC biological response to hypoxia conditions, or hypoxia-associated  
 CC pathology in a patient. The coding sequences and proteins are also useful  
 CC for monitoring the therapeutic treatment of a disease or physiological  
 CC condition, such as cancer, ischaemic conditions, reperfusion injury,  
 CC retinopathy, neonatal stress, preclampsia, atherosclerosis, inflammatory  
 CC conditions, wound healing, inflammation, erythropoiesis or hair loss

XX Sequence 215 AA;

Query Match 38.6%; Score 22; DB 5; Length 215;  
 Best Local Similarity 36.4%; Pred. No. 5e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYF 29  
Db 90 FTASKNGTYKF 100

RESULT 8  
ABBI1780  
ID ABBI1780 standard; peptide: 219 AA.  
XX  
AC ABBI1780;  
XX  
DT 11-JAN-2002 (first entry)  
XX  
DE Human T1 receptor ligand III homologue, SEQ ID NO:2150.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; antiinflammatory;  
KW antidiabetic; antidiabetic; haemostatic; antiarteriosclerotic;  
KW cytoskeletal; osteopathic; vasotropic; cardiac; virucide; antibacterial;  
KW antifungal; vulnery; antitumor.

XX  
OS Homo sapiens.  
XX  
PN W0200157188-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US003800.  
XX  
PR 03-FEB-2000; 2000US-00496914.  
XX  
PR 27-APR-2000; 2000US-00560875.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR N-PSDB; ABA09024.  
XX  
PT Human proteins and DNA encoding sequences useful for preventing, treating  
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis  
XX  
PS Claim 20; Page 245; 1963pp; English.

XX  
CC Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical

CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness.  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a novel human  
CC polypeptide of the invention

XX  
SQ Sequence 219 AA:  
XX

Qy 19 FXXXXXGXYF 29  
Db 93 FTASKNGTYKF 103

Query Match 38.6%; Score 22; DB 4; Length 219;  
Best Local Similarity 36.4%; Pred. No. 5.1e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 9  
AAM93844  
ID AAM93844 standard; protein: 224 AA.  
XX  
AC AAM93844;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide, SEQ ID NO: 3923.  
XX  
KW Human, full length cDNA; cDNA synthesis; oligo-capping.  
XX  
OS Homo sapiens.  
XX  
PN EP1130094-A2.  
XX  
PD 05-SEP-2001.  
XX  
PF 07-JUL-2000; 2000EP-00114089.  
XX  
PR 08-JUL-1999; 99JP-00194486.  
XX  
PR 11-JAN-2000; 2000JP-00118774.  
XX  
PR 02-MAY-2000; 2000JP-00183765.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX  
DR N-PSDB; AAK94800.  
XX  
PT 830 Primers useful for synthesizing full length cDNA clones and their use  
PT in genetic manipulation.  
XX  
PS Claim 8; SEQ ID NO 3923; 1380pp + Sequence Listing; English.

XX  
CC The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been isolated  
CC and nucleotide sequences of 5' - and 3' - ends of the cDNA molecules have  
CC been determined. Primers for synthesizing the full length cDNA are useful  
CC for clarifying the function of the protein encoded by the cDNA. The full  
CC length clones were obtained by construction of full length enriched cDNA

CC libraries that were synthesised by the oligo-capping method. The primers  
CC enable the production of the full length cDNA easily without any special  
CC methods. The present sequence is a polypeptide encoded by a full length  
CC human cDNA of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in CD-ROM  
CC format directly from EPO

XX Sequence 224 AA;

Query Match 38.6%; Score 22; DB 4; Length 224;

Best Local Similarity 36.4%; Pred. No. 5.3e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

19 FXXXXXGXYP 29

98 FTASKNGTYKF 108

RESULT 10

AAB92820 standard; protein; 224 AA.

AC AAB92820;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:11348.

KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-00116126.

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-

PT length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length

PT cDNAs.

PS Claim 8; SEQ ID NO 11348; 2537bp + Sequence listing; English.

XX The present invention describes primer sets for synthesising 5602 full-

XX length cDNAs defined in the specification. Where a primer set comprises:

XX (a) an oligo-dT primer and an oligonucleotide complementary to the

XX complementary strand of a polynucleotide which comprises one of the 5602

XX nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in the

XX specification. The primer sets can be used in antisense therapy and in

XX gene therapy. The primers are useful for synthesising polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893

CC represent human amino acid sequences; and AAH13629 to AAH1632 represent

CC oligonucleotides, all of which are used in the exemplification of the

CC present invention

XX Sequence 224 AA;

Query Match 38.6%; Score 22; DB 4; Length 224;

Best Local Similarity 36.4%; Pred. No. 5.3e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

19 FXXXXXGXYP 29

98 FTASKNGTYKF 108

RESULT 11

ADL1890

ID ADL1890 standard; protein; 224 AA.

AC ADL1890;

DT 20-MAY-2004 (first entry)

DE Human protein encoded by a full length cDNA clone SegID 3923.

KM human; medicine; signal transduction; glycoprotein; transcription;

OS oligo-capping method.

OS Homo sapiens.

PN EP1396543-A2.

PD 10-MAR-2004.

PF 07-JUL-2000; 2003EP-00025638.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183865.

PR 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2004-204755/20.

XX N-PSDB; ADL1889.

PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full

PT length human cDNAs.

PS Example 1; SEQ ID NO 3923; 1340bp; English.

XX This invention relates to a novel primers useful for synthesising full

XX length cDNA molecules that encode human proteins. Specifically, it refers

XX to secretory or membrane proteins that are potential therapeutic agents/

XX target molecules in the field of medicine, and in particular genes

XX encoding proteins that are associated with signal transduction.

XX glycoproteins and transcription. The present invention describes a method

XX for efficiently cloning a full length human cDNA from both the 5' and 3'

XX ends using the oligo-capping method. This polypeptide sequence is a full

XX length human protein of the invention.

XX Sequence 224 AA;

Query Match 38.6%; Score 22; DB 8; Length 224;

Best Local Similarity 36.4%; Pred. No. 5.3e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGYXFP 29  
 Db 98 FTASKNGTYKF 108

RESULT 12  
 ID ABR41674 standard; protein; 231 AA.  
 XX  
 AC ABR41674;  
 XX  
 DT 02-JUN-2003 (first entry)  
 XX  
 DE Human DITHP cell membrane protein.  
 XX  
 KM Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;  
 KM cancer; cell proliferative disorder; autoimmune disorder;  
 KM inflammatory disorder; infection; hormonal disorder; metabolic disorder;  
 KM neurological disorder; gastrointestinal disorder; transport disorder;  
 KM connective tissue disorder; drug screening; proteome analysis;  
 KM gene therapy; antisense therapy; genotyping; transgenic animal; knock in;  
 KM disease model; toxicological testing; transcript imaging;  
 KM cell membrane protein.  
 XX  
 OS Homo sapiens.  
 PN WO200297031-A2.  
 PD 05-DEC-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-US010056.  
 XX  
 PR 28-MAR-2001; 2001US-0279619P.  
 PR 29-MAR-2001; 2001US-0280067P.  
 PR 16-MAR-2001; 2001US-0280068P.  
 PR 16-MAY-2001; 2001US-0291280P.  
 PR 17-MAY-2001; 2001US-0291829P.  
 PR 17-MAY-2001; 2001US-0291849P.  
 PR 19-JUN-2001; 2001US-0299428P.  
 PR 20-JUN-2001; 2001US-0299776P.  
 PR 20-JUN-2001; 2001US-0300001P.  
 XX  
 PA (INCY-) INCTRE GENOMICS INC.  
 XX  
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshery SR;  
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;  
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 PI Flores V, Marwaha R, Lo A, Lan RY, Uraahka ME;  
 XX  
 DR WPI; 2003-129518/12.  
 DR N-PSDB; ACC46611.  
 XX  
 PT Novel human diagnostic and therapeutic polypeptide useful for identifying  
 PT test compound which specifically binds to a polypeptide encoded by human  
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.  
 XX  
 PS Claim 27; SEQ ID NO 1209; 591pp; English.  
 XX  
 CC The invention relates to novel human diagnostic and therapeutic  
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded  
 CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to  
 CC polynucleotide sequences at least 90% identical to the dithp cDNA  
 CC sequences of the invention; recombinant vectors, host cells and  
 CC transgenic organisms comprising a dithp nucleic acid sequence; the  
 CC recombinant production of DITHP proteins; antibodies specific for DITHP  
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods of  
 CC detecting dithp nucleotide and protein sequences; methods of screening  
 CC for compounds which specifically bind a DITHP protein; and methods of  
 CC assessing the toxicity of test compounds using a dithp hybridisation  
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the  
 CC diagnosis of a wide variety of conditions including cancer and other cell

CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,  
 CC viral, fungal or parasitic infections; hormonal disorders; metabolic  
 CC disorders; neurological disorders; gastrointestinal disorders; transport  
 CC disorders; and connective tissue disorders. They may also be used to  
 CC screen for modulators of protein activity or gene expression. DITHP  
 CC proteins can additionally be used in analysis of the proteome of a tissue  
 CC or cell type and to induce antibodies. The dithp nucleic acids are  
 CC additionally useful in somatic or germline gene therapy of the disorders  
 CC mentioned above, as a source of antisense sequences, as a source of  
 CC probes and primers, in genotyping and identification of individuals, in  
 CC the generation of transgenic animal models of human disease or knock in  
 CC humanised animals, in toxicological testing, and in transcript imaging.  
 CC The present sequence represents a DITHP protein which is a cell membrane  
 CC protein. Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SQ Sequence 231 AA;  
 Qy 19 FXXXXXGYXFP 29  
 Db 105 FTASKNGTYKF 115

Query Match 38.6%; Score 22; DB 6; Length 231;  
 Best Local Similarity 36.4%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 13  
 ID AAG82486 standard; protein; 266 AA.  
 XX  
 AC AAG82486;  
 XX  
 DT 03-SEP-2001 (first entry)  
 XX  
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:2066.  
 XX  
 KM Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
 KM endocarditis.  
 XX  
 OS Staphylococcus epidermidis.  
 PN WO200134809-A2.  
 PD 17-MAY-2001.  
 XX  
 PR 09-NOV-2000; 2000WO-US030782.  
 PR 09-NOV-1999; 99US-0164258P.  
 XX  
 PA (GLAXO) GLAXO GROUP LTD.  
 XX  
 PI Kimmberly WJ;  
 XX  
 DR WPI; 2001-316495/33.  
 DR N-PSDB; AAH53336.  
 XX  
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
 PT useful for vaccinating against infections, e.g. endocarditis.  
 XX  
 PS Claim 18; Page 560; 2188pp; English.  
 XX  
 CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)  
 CC and (II) can have antibacterial activity and therefore can be used in  
 CC vaccination. The nucleic acids (I) may be used to produce the S.  
 CC epidermidis polypeptides (II) via the production of vectors containing  
 CC them which are used to produce host cells which express the  
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
 CC used to vaccinate subjects and to raise antibodies against the bacteria.  
 CC The polypeptides may also be used to assay for other inhibitors of their  
 CC activity and therefore identify compounds that may be used for the

CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to  
CC AAH55090 represent specifically claimed *S. epidermidis* genomic DNA  
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098  
CC represent oligonucleotide sequences and primers which are used in the  
CC exemplification of the present invention. N.B. The present invention  
CC specifically claims all the polynucleotide sequences given in the  
CC sequence listing of the present specification, however the sequence  
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
CC for SEQ ID NO:4455 to 4464  
XX  
SQ Sequence 266 AA;  
XX  
QY Query Match 38.6%; Score 22; DB 4; Length 266;  
Best Local Similarity 36.4%; Pred. No. 6.2e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
19 FXXXXXGXXF 29  
147 FKTRSGVYTF 157  
Db  
XX  
RESULT 14  
ABO69527  
ID ABO69527 standard; protein; 268 AA.  
XX  
AC ABO69527;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #1702.  
XX  
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
DR WPI; 2003-615309/58.  
DR N-PSDB; ABD03098.  
XX  
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 18273; 455pp; English.  
XX  
CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at

CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 268 AA;  
XX  
QY Query Match 38.6%; Score 22; DB 7; Length 268;  
Best Local Similarity 36.4%; Pred. No. 6.3e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
19 FXXXXXGXXF 29  
41 FLAVATGTYAF 51  
Db  
XX  
RESULT 15  
ABP39297  
ID ABP39297 standard; protein; 292 AA.  
XX  
AC ABP39297;  
XX  
DT 24-JUL-2002 (first entry)  
XX  
DE Staphylococcus epidermidis ORF amino acid sequence SRQ ID NO:4142.  
XX  
KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
KM antibacterial; gene therapy.  
XX  
OS Staphylococcus epidermidis.  
XX  
PN US6380370-B1.  
XX  
PD 30-APR-2002.  
XX  
PF 13-AUG-1998; 98US-00134001.  
XX  
PR 14-AUG-1997; 97US-0055779P.  
PR 08-NOV-1997; 97US-0064964P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Doucette-Stamm LA, Bush D;  
XX  
DR WPI; 2002-381255/41.  
DR N-PSDB; ABN91842.  
XX  
PT Novel isolated nucleic acid encoding a Staphylococcus epidermis  
PT polypeptide, useful for diagnosing and treating bacterial infections.  
XX  
PS Disclosure; SEQ ID NO 4142; 267pp; English.  
XX  
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
CC given in ABP35124 to ABP37960. The *S. epidermidis* sequences have  
CC antibacterial activity and can be used in gene therapy. The sequences can  
CC also be used in the diagnosis and treatment of bacterial infections,  
CC particularly *S. epidermidis* infections. The sequences can be used to  
CC screen for compounds able to interfere with the *S. epidermidis* life cycle  
CC or inhibit *S. epidermidis* infection. N.B. The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the USPTO web site  
XX  
SQ Sequence 292 AA;  
XX  
QY Query Match 38.6%; Score 22; DB 5; Length 292;  
Best Local Similarity 36.4%; Pred. No. 6.9e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
19 FXXXXXGXXF 29  
173 FKTRSGVYTF 183  
Db  
XX  
Search completed: December 25, 2004, 08:30:35  
Job time : 49.6487 secs



: Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:08:11 : Search time 7.9462 Seconds  
(without alignments) 375.364 Million cell updates/sec

Title: US-10-621-787-5  
Perfect score: 57  
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXXGXGXFX 31

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 79:\*

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	38.6	250	2	E83651
2	22	38.6	302	2	D82112
3	22	38.6	421	2	A99309
4	22	38.6	423	2	H90490
5	22	38.6	441	2	T19306
6	22	38.6	868	2	T02635
7	22	38.6	3705	2	AD0123
8	22	38.6	4976	2	T14165
9	21	36.8	156	2	E72052
10	21	36.8	156	2	A86573
11	21	36.8	210	2	F64609
12	21	36.8	210	2	B71906
13	21	36.8	212	2	H81344
14	21	36.8	231	2	B83060
15	21	36.8	238	2	I64038
16	21	36.8	269	2	T49293
17	21	36.8	277	2	T06760
18	21	36.8	303	2	T35616
19	21	36.8	303	2	A12543
20	21	36.8	307	2	A71602
21	21	36.8	325	2	T31989
22	21	36.8	420	2	AG3057
23	21	36.8	420	2	G98238
24	21	36.8	450	2	B96561
25	21	36.8	508	2	T04605
26	21	36.8	516	2	C82946
27	21	36.8	555	2	T01391
28	21	36.8	653	2	T03319
29	21	36.8	669	2	T24571

30	21	36.8	704	2	T13665	NADH2 dehydrogenas
31	21	36.8	724	2	AB0531	lysine decarboxyla
32	21	36.8	724	2	A42953	nitrous oxide redu
33	21	36.8	724	2	D96949	secreted protein c
34	21	36.8	779	2	T44659	nitrous oxide redu
35	21	36.8	826	2	E81706	conserved hypochet
36	21	36.8	885	2	AC2444	hypothetical prote
37	21	36.8	976	2	G82209	GDDR family prote
38	21	36.8	1138	2	G71554	probable transmemb
39	21	36.8	1316	2	B86240	protein F20B24.12
40	21	36.8	1417	2	H90670	probable invasin l
41	21	36.8	1417	2	D85521	probable adhesin e
42	21	35.1	40	2	G81731	hypothetical prote
43	20	35.1	127	2	A53257	H+-transporting AT
44	20	35.1	127	2	H64012	suft protein homol
45	20	35.1	129	2	B90175	NADH dehydrogenase
46	20	35.1	171	2	C81280	probable periplasm
47	20	35.1	173	2	I39569	hypothetical prote
48	20	35.1	181	2	G75071	phenylacrylic acid
49	20	35.1	181	2	A71094	probable 3-octapre
50	20	35.1	187	2	AF2009	hypothetical prote
51	20	35.1	219	2	F97083	protein of phospho
52	20	35.1	240	2	H69535	acetylglutamate ki
53	20	35.1	242	2	I78931	arylamine N-acetyl
54	20	35.1	246	2	T13747	hypothetical prote
55	20	35.1	260	2	D24584	spectinomycin resi
56	20	35.1	267	2	B69752	beta-lactamase hom
57	20	35.1	282	2	G72369	hemk protein - The
58	20	35.1	290	2	I78930	arylamine N-acetyl
59	20	35.1	290	2	B61267	arylamine N-acetyl
60	20	35.1	290	2	I58425	lipopolyasaccharide
61	20	35.1	297	2	F64306	oligopeptide trans
62	20	35.1	298	2	AC3617	hypothetical prote
63	20	35.1	299	2	H64354	conserved hypochet
64	20	35.1	299	2	AH0367	glutamate decarbox
65	20	35.1	304	2	AE3623	conserved hypochet
66	20	35.1	313	2	A95859	protein C50H11.6 (
67	20	35.1	317	2	F88986	gluconolactonase (
68	20	35.1	320	2	S28218	hypothetical prote
69	20	35.1	323	2	T25459	multidrug resistanc
70	20	35.1	334	2	T19860	hypothetical prote
71	20	35.1	353	2	A88987	protein C50H11.4 (
72	20	35.1	364	2	T43361	probable potassium
73	20	35.1	385	2	G64251	hypothetical prote
74	20	35.1	385	2	S73488	hypothetical prote
75	20	35.1	417	2	D90304	amino acid transpo
76	20	35.1	418	2	B75087	glutamate-1-semial
77	20	35.1	424	2	E81358	glutamate-1-semial
78	20	35.1	426	2	AG0411	hypothetical prote
79	20	35.1	437	2	T32763	hypothetical prote
80	20	35.1	438	2	A57720	kinase receptor a
81	20	35.1	462	2	G95233	Cof family protein
82	20	35.1	462	2	A86717	conserved hypochet
83	20	35.1	464	2	E83029	replicative DNA he
84	20	35.1	465	2	C81751	probable sodium-tr
85	20	35.1	465	2	F71489	sodium-translocati
86	20	35.1	466	1	B43332	glutamate decarbox
87	20	35.1	466	1	S24234	glutamate decarbox
88	20	35.1	466	1	E91178	glutamate decarbox
89	20	35.1	466	2	F86024	glutamate decarbox
90	20	35.1	466	2	G85726	glutamate decarbox
91	20	35.1	466	2	B90891	glutamate decarbox
92	20	35.1	467	2	D86583	ubiquinone oxidore
93	20	35.1	467	2	E72040	probable sodium-tr
94	20	35.1	472	2	A99098	conserved hypochet
95	20	35.1	498	1	A60028	nucleoprotein - in
96	20	35.1	498	1	VHIV34	nucleoprotein - in
97	20	35.1	498	1	VHIV34	nucleoprotein - in
98	20	35.1	498	1	VHIV34	nucleoprotein - in
99	20	35.1	498	1	VHIV34	nucleoprotein - in
100	20	35.1	498	1	VHIV34	nucleoprotein - in

## ALIGNMENTS

## RESULT 1

E83651 hypothetical protein BH0013 [imported] - Bacillus halodurans (strain C-125)

C/Species: Bacillus halodurans

C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004

C/Accession: E83651

R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saseaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A/Reference number: A83650; MUID:20512582; PMID:11058132

A/Accession: E83651

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-250 &lt;STO&gt;

A/Cross-references: UNIPROT:Q9KGP4; GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BAB037

A/Experimental source: strain C-125

C/Genetics:

A/Gene: BH0013

## Query Match

Best Local Similarity 38.6%; Score 22; DB 2; Length 250;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 132 FTFSGAGSYDF 142

## RESULT 2

D82112 tyra protein VC2145 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C/Species: Vibrio cholerae

C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C/Accession: D82112

R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: D82112

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-302 &lt;HEI&gt;

A/Cross-references: UNIPROT:Q9KQ59; GB:AE004287; GB:AE003852; NID:g9656689; PIDN:AAF9528

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

C/Genetics:

A/Gene: VC2145

A/Map position: 1

C/Superfamily: Escherichia coli hypothetical protein b2431

## Query Match

Best Local Similarity 38.6%; Score 22; DB 2; Length 302;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 277 FTAAVTAATVF 287

## RESULT 3

A99309 membrane transporter SSO1505 [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C/Accession: A99309

R/She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.

arrett, R.A.; Ragan, M.A.; Senen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: A99309

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-421 &lt;KUR&gt;

A/Cross-references: UNIPROT:Q97Y43; GB:AB006641; NID:g13814734; PIDN:AAK41728.1; GSPDB:G

C/Genetics:

A/Gene: SSO1505

## Query Match

Best Local Similarity 38.6%; Score 22; DB 2; Length 421;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 338 FAATAATGYSF 348

## RESULT 4

H90490 membrane transporter SSO3079 [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C/Accession: H90490

R/She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.

arrett, R.A.; Ragan, M.A.; Senen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: H90490

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-423 &lt;KUR&gt;

A/Cross-references: UNIPROT:Q97UD8; GB:AB006641; NID:g13816494; PIDN:AAK43183.1; GSPDB:G

C/Genetics:

A/Gene: SSO3079

## Query Match

Best Local Similarity 38.6%; Score 22; DB 2; Length 423;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 340 FAATAATGYSF 350

## RESULT 5

T19306 hypothetical protein C15C8.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T19306

R/Sims, M.

submitted to the EMBL Data Library, June 1996

A/Reference number: Z19106

A/Accession: T19306

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-441 &lt;MTL&gt;

A/Cross-references: UNIPROT:Q18017; EMBL:Z75527; PIDN:CAA99774.1; GSPDB:GN00023; CESP:CI;

A/Experimental source: clone C15C8

C/Genetics:

A/Gene: CESP:C15C8.1

A/Map position: 5

A/Intons: 156/2; 205/3; 258/1

## Query Match

Best Local Similarity 38.6%; Score 22; DB 2; Length 441;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

Db 165 PAALISGSYHF 175

## RESULT 6

D2 protein homolog - slime mold (Dictyostelium discoideum) plasmid Dp5  
C:Species: Dictyostelium discoideum  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C:Accession: T02635  
R:Rieben J.F., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyosawa, H.; Hugh  
Genetics 148, 1117-1125, 1998  
A:Title: Dictyostelium discoideum nuclear plasmid Dp5 is a chimera related to the Dp1  
A:Reference number: Z14684; MUID:98198836; PMID:9539429  
A:Accession: T02635  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-868 <RIB>  
A:Cross-references: UNIPROT:O60984; EMBL:AF000580; NID:g3068582; PIDN:AA014391.1; PID:g3  
A:Experimental source: strain WS2162  
C:Genetics:  
A:Gene: d2  
A:Genome: plasmid  
A:Mobile element: plasmid Dp5  
C:Superfamily: slime mold (Dictyostelium discoideum) plasmid Dp5 D2 protein homolog

Query Match 38.6%; Score 22; DB 2; Length 868;  
Best Local Similarity 36.4%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYXF 29

Db 345 FTTYSKGRYSP 355

## RESULT 7

AD0123  
probable autotransporter protein yaph [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AD0123  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
demo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AD0123  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-3705 <KUR>  
A:Cross-references: UNIPROT:Q8ZHA1; GB:AL590842; PIDN:CAC9847.1; PID:g15979073; GSPDB:C  
C:Genetics:  
A:Gene: yaph

Query Match 38.6%; Score 22; DB 2; Length 3705;  
Best Local Similarity 36.4%; Pred. No. 3.3e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYXF 29

Db 3553 FTASVEGGYAF 3563

## RESULT 8

T14165  
peptide synthetase homolog - Mycobacterium smegmatis  
C:Species: Mycobacterium smegmatis  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: T14165  
R:Yu, S.; Flier, E.; Jacobs Jr., W.R.  
J. Bacteriol. 180, 4676-4685, 1998  
A:Title: Analysis of the exochelin locus in mycobacterium smegmatis: biosynthesis genes

A:Reference number: Z17898; MUID:98389687; PMID:9721311

A:Accession: T14165  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4976 <YUS>

A:Cross-references: UNIPROT:O87314; EMBL:AF027770; NID:g3560502; PID:g3560507; PIDN:AA08  
C:Genetics:  
A:Gene: fxbC

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein  
F:68-566/Domain: acetate-CoA ligase homology <ACL>  
F:582-649/Domain: acyl carrier protein homology <ACP1>  
F:1620-2062/Domain: acetate-CoA ligase homology <ACLP1>  
F:2078-2144/Domain: acyl carrier protein homology <ACP2>  
F:3139-3591/Domain: acetate-CoA ligase homology <ACL2>  
F:3608-3679/Domain: acyl carrier protein homology <ACP3>  
F:4198-4620/Domain: acetate-CoA ligase homology <ACL3>  
F:4637-4705/Domain: acyl carrier protein homology <ACP4>  
F:614,3643,4669/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 38.6%; Score 22; DB 2; Length 4976;  
Best Local Similarity 36.4%; Pred. No. 4.5e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYXF 29

Db 1738 FAGTSGLYDP 1748

## RESULT 9

E72052  
RNA methylase - Chlamydia pneumoniae (strain CWL029)  
C:Species: Chlamydia pneumoniae  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: E72052  
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lamell, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388  
A:Accession: E72052  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-156 <ARN>  
A:Cross-references: UNIPROT:Q9Z7P4; GB:AE001648; GB:AE001363; NID:g4376946; PIDN:AA018791;  
A:Experimental source: strain CWL029  
C:Genetics:  
A:Gene: spoU\_2  
C:Superfamily: Chlamydia pneumoniae RNA methylase

Query Match 36.8%; Score 21; DB 2; Length 156;  
Best Local Similarity 36.4%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYXF 29

Db 88 FSLPSSGTYVF 98

## RESULT 10

A86573  
RNA methylase [imported] - Chlamydia pneumoniae (strain J138)  
C:Species: Chlamydia pneumoniae  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: A86573  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349; PMID:10871362  
A:Accession: A86573  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-156 <STO>  
A:Cross-references: UNIPROT:Q9Z7P4; GB:BA000008; NID:g8979032; PIDN:BA08867.1; GSPDB:GN1  
A:Experimental source: strain J138

C/Genetics:  
A/Gene: spou 2  
C/Superfamily: Chlamydomonada pneumoniae rRNA methylase

Query Match 36.8%; Score 21; DB 2; Length 156;  
Best Local Similarity 36.4%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29  
DB 88 FSLPSSGTYVF 98

RESULT 11  
F64609  
conserved hypothetical integral membrane protein HP0718 - Helicobacter pylori (strain 26)

C/Species: Helicobacter pylori  
C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C/Accession: F64609

R/Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.; Nature 388, 539-547, 1997  
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A/Reference number: A64520; MUID:97394467; PMID:9252185  
A/Accession: F64609

A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-210 <TOM>  
A/Cross-references: UNIPROT:O25420; GB:AE000585; GB:AE000511; NID:92313845; PIDN:AA00777

C/Superfamily: Escherichia coli hypothetical 23k protein (sbm-fba intergenic region)

Query Match 36.8%; Score 21; DB 2; Length 210;  
Best Local Similarity 36.4%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29  
DB 75 FGALFTGYFVF 85

RESULT 12  
B71906

C/Superfamily: Helicobacter pylori (strain J99)

C/Species: Helicobacter pylori

A/Variety: strain J99  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

C/Accession: B71906

R/Alt, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.  
A/Reference number: A71800; MUID:99120557; PMID:9923682  
A/Accession: B71906

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-210 <ARN>  
A/Cross-references: UNIPROT:Q9ZIC4; GB:AE001497; GB:AE001439; NID:94155199; PIDN:AA00622

A/Experimental source: strain J99

C/Genetics:

A/Gene: jhp0656  
C/Superfamily: Escherichia coli hypothetical 23k protein (sbm-fba intergenic region)

Query Match 36.8%; Score 21; DB 2; Length 210;  
Best Local Similarity 36.4%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29  
DB 75 FGALFTGYFVF 85

RESULT 13  
H81344

C/Superfamily: Campylobacter jejuni (strain NCTC 11168)

C/Species: Campylobacter jejuni

C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004

C/Accession: H81344

R/Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrett, N.; Nature 403, 665-668, 2000  
A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A/Reference number: A81250; MUID:20150912; PMID:10688204

A/Accession: H81344

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-212 <PAR>

A/Cross-references: UNIPROT:Q9PPH4; GB:AL139076; GB:AL111168; NID:96968128; PIDN:CA87300;

A/Experimental source: serotype O2, strain NCTC 11168

C/Genetics:

A/Gene: Cj0733

Query Match 36.8%; Score 21; DB 2; Length 212;  
Best Local Similarity 36.4%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29  
DB 137 FSGKAVGYVSF 147

RESULT 14  
B83060

C/Superfamily: Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C/Accession: B83060

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.; Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.  
A/Reference number: A82950; MUID:20437337; PMID:10984043  
A/Accession: B83060

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-231 <STO>

A/Cross-references: UNIPROT:Q9HVB0; GB:AE004882; GB:AE004091; NID:9950939; PIDN:AA00807;

A/Experimental source: strain PA01

C/Genetics:

A/Gene: PA4685

Query Match 36.8%; Score 21; DB 2; Length 231;  
Best Local Similarity 36.4%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29  
DB 146 FASEDSGVYVF 156

RESULT 15  
I64038

C/Superfamily: Haemophilus influenzae (strain Rd KW20)

C/Species: Haemophilus influenzae

C/Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004

C/Accession: I64038

R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kervatage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995  
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A/Reference number: A64000; MUID:95350630; PMID:7542800  
A/Accession: I64038

A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-238 <TTGR>  
A:Cross-references: UNIPROT:P44278; GB:U32835; GB:I42023; NID:g1574459; PIDN:AAC23269.1

Query Match 36.8%; Score 21; DB 2; Length 238;  
Best local similarity 36.4%; Pred. No. 3.7e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGYYX 29  
DB 122 PALAGMGSYF 132

Search completed: December 25, 2004, 08:39:40  
Job time : 12.9462 secs

This Page Blank (upto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 25, 2004, 06:42:17 ; Search time 46.0095 Seconds

(without alignments)  
387.672 Million cell updates/sec

Title: US-10-621-787-5

Perfect score: 57  
Sequence: 1 PXXXXXXXXXXXXXXXXXXXXGXXFXFX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 100 summaries

Database : UniProt 02:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	42.1	1020	2 Q89DX5	Q89dx5 bradyrhizob
2	23	40.4	214	2 Q7RAT5	Q7rat5 plasmidium
3	23	40.4	352	2 Q9DED3	Q9ded3 oncorhynch
4	23	40.4	483	1 PRTA_PHOAZ	P82115 photorhabd
5	23	40.4	483	2 Q84F72	Q84f72 photorhabd
6	23	40.4	483	2 AAS19409	Aas19409 photorhab
7	23	40.4	618	2 Q92LIN9	Q92lin9 rhizobium m
8	22	38.6	144	2 Q8BRP1	Q8brp1 mus musculu
9	22	38.6	180	2 Q6N9E2	Q6n9e2 rhodospheuo
10	22	38.6	180	2 CAE27048	Caet27048 rhodosphe
11	22	38.6	188	2 Q8NBU8	Q8nbu8 homo sapien
12	22	38.6	215	1 YAO9_HUMAN	Q9y3b3 homo sapien
13	22	38.6	217	2 Q6NMW3	Q6nmw3 brachydanio
14	22	38.6	217	2 AAH67383	Aah67383 brachydan
15	22	38.6	219	2 Q6P2V7	Q6p2v7 xenopus tro
16	22	38.6	219	2 AAH64279	Aah64279 xenopus t
17	22	38.6	224	2 Q96K51	Q96k51 homo sapien
18	22	38.6	224	2 Q8WU6	Q8wu6 homo sapien
19	22	38.6	229	2 Q848M1	Q848m1 uncultured
20	22	38.6	250	2 Q9KGP4	Q9kgp4 bacillus ha
21	22	38.6	272	2 Q6PWJ1	Q6pwj1 ralsstonia s
22	22	38.6	272	2 AAS87586	Aas87586 ralsstonia
23	22	38.6	290	2 Q8CSOS	Q8csos staphylococ
24	22	38.6	296	2 Q9CNH6	Q9cnh6 pasteurella
25	22	38.6	298	2 Q6LTW2	Q6ltw2 photobacter
26	22	38.6	298	2 QAG19273	Qag19273 photobact
27	22	38.6	299	2 Q6D8U0	Q6d8u0 erwinia car
28	22	38.6	300	2 Q7MIM2	Q7mim2 vibrio vuln
29	22	38.6	300	2 Q87MJ3	Q87mj3 vibrio para
30	22	38.6	300	2 Q8DBA0	Q8dba0 vibrio vuln
31	22	38.6	302	2 Q9KQ59	Q9kq59 vibrio chol

32	22	38.6	323	2 Q8PXN5	Q8pxn5 methanosarc
33	22	38.6	325	2 Q8PSW8	Q8psw8 methanosarc
34	22	38.6	355	2 Q966G5	Q966g5 caenorhabdi
35	22	38.6	384	2 Q8EJ24	Q8ej24 shewanella
36	22	38.6	398	2 Q744P3	Q744p3 mycobacteri
37	22	38.6	398	2 AAS02705	Aas02705 mycobacte
38	22	38.6	404	2 Q8ULU7	Q8ulu7 pyrococcus
39	22	38.6	404	2 Q6UTJ2	Q6utj2 homo sapien
40	22	38.6	404	2 AAQ97430	Aaq97430 homo sapi
41	22	38.6	404	2 AAQ97433	Aaq97433 homo sapi
42	22	38.6	409	1 RUN3_MOUSE	Q64131 m runt-reia
43	22	38.6	409	2 Q91ZK1	Q91zk1 ralsstonia s
44	22	38.6	421	2 Q97Y43	Q97y43 sulfolobus
45	22	38.6	422	2 Q7ODI2	Q7odi2 anopheles g
46	22	38.6	423	2 Q97UD8	Q97ud8 sulfolobus
47	22	38.6	440	2 Q8XPA4	Q8xpa4 clostridium
48	22	38.6	441	2 Q18017	Q18017 photorhabd
49	22	38.6	480	1 PRTA_PHOLU	Q84f70 photorhabd
50	22	38.6	486	1 PRTA_PHOLU	Q7g448 anopheles g
51	22	38.6	519	2 Q7O4F8	Q7o4f8 anopheles g
52	22	38.6	519	2 Q86GR1	Q86gr1 anopheles g
53	22	38.6	529	2 Q9FGF7	Q9fgf7 arabidopsis
54	22	38.6	608	2 Q7OSX0	Q7osx0 giardia lam
55	22	38.6	661	2 Q7QW29	Q7qw29 giardia lam
56	22	38.6	679	2 Q8XW21	Q8xw21 ralsstonia s
57	22	38.6	743	2 Q89WU9	Q89wu9 bradyrhizob
58	22	38.6	763	2 Q6KGV2	Q6kgv2 bacterioph
59	22	38.6	787	2 Q9GSD4	Q9gsd4 plasmidium
60	22	38.6	859	2 Q7PDL4	Q7pdl4 plasmidium
61	22	38.6	868	2 Q60984	Q60984 dictyosteli
62	22	38.6	973	2 Q7PJ63	Q7pj63 anopheles g
63	22	38.6	1098	2 Q8A049	Q8a049 bacteroides
64	22	38.6	1127	2 Q8L3A4	Q8l3a4 bacillus st
65	22	38.6	1457	2 Q9F289	Q9f289 yersinia pe
66	22	38.6	1677	2 Q9AS94	Q9as94 oryza sativ
67	22	38.6	1913	2 Q9GRJ6	Q9grj6 leucophaea
68	22	38.6	2215	2 Q7WBN0	Q7wb0 bordetella
69	22	38.6	3346	2 Q7WNS4	Q7wns4 bordetella
70	22	38.6	3705	2 Q9F285	Q9f285 yersinia pe
71	22	38.6	3705	2 Q8ZHA1	Q8zha1 yersinia pe
72	22	38.6	3710	2 Q74QP7	Q74qp7 yersinia pe
73	22	38.6	3710	2 Q8CZU2	Q8czu2 yersinia pe
74	22	38.6	3710	2 AAS63576	Aas63576 yersinia
75	22	38.6	4976	2 Q87314	Q87314 mycobacteri
76	22	38.6	5020	2 Q8B9W3	Q8b9w3 shewanella
77	22	38.6	6939	2 Q7RIS4	Q7rie4 giardia lam
78	22	36.8	45	2 Q8F630	Q8f630 leptospira
79	21	36.8	109	2 Q9NHR7	Q9nh7 plasmidium
80	21	36.8	109	2 Q9NHL5	Q9nh5 plasmidium
81	21	36.8	113	2 Q86TX2	Q86tx2 dictyosteli
82	21	36.8	125	2 Q6MLU8	Q6mlu8 bdellovibri
83	21	36.8	125	2 CAE79758	Caet79758 bdellovib
84	21	36.8	135	2 Q8FAV7	Q8fa7 escherichia
85	21	36.8	136	2 Q6KGS1	Q6kgs1 macaca mula
86	21	36.8	136	2 AAQ02794	Aaq02794 macaca mu
87	21	36.8	139	2 Q7Z239	Q7z239 plasmidium
88	21	36.8	154	2 Q9L5R0	Q9l5r0 salmoneella
89	21	36.8	154	2 Q935R2	Q935r2 salmoneella
90	21	36.8	156	2 Q9Z7P4	Q9z7p4 chlamydia p
91	21	36.8	159	2 Q8C129	Q8c129 mus musculu
92	21	36.8	169	2 Q8E085	Q8e085 streptococc
93	21	36.8	169	2 Q8ESW2	Q8esw2 streptococc
94	21	36.8	175	1 Q995_THETN	Q995t thermotacarc
95	21	36.8	187	1 Y177_METWA	Q9q019 methanosarc
96	21	36.8	192	1 Q16605	Q16605 caenorhabdi
97	21	36.8	195	1 YG14_METAC	Q8q60 methanosarc
98	21	36.8	196	1 P24_CRIGR	Q9j020 cricetulum
99	21	36.8	196	1 YB05_THRAC	Q9h167 thermoplasm
100	21	36.8	201	1 P24_HUMAN	Q15363 homo sapien

## ALIGNMENTS

```

RESULT 1
ID 089DX5 PRELIMINARY; PRT; 1020 AA.
AC 089DX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AcbB/AcrB/AcrP family protein.
OS OrderedLocustNames=bl17312;
OC Bradyrhizobium japonicum.
OC Bradyrhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USD110;
RA MEDLINE=22484998; PubMed=12597275;
RA Kaneke T., Nakamura Y., Sato S., Minamideawa K., Uchiyumi T.,
RA Sasanoto S., Watanabe A., Idegawa K., Irituchi M., Kawashima K.,
RA Kobara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USD110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005961; BAC52577.1; -.
DR HSSP; P31224; 11WG.
DR InterPro; IPR001036; Acrlflvin_res.
DR Pfam; PF00873; ACR_tran; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
KM Complete proteome.
SQ SEQUENCE 1020 AA; 109742 MW; F593D2PF5D6FDB9D CRC64;

Query Match 42.1%; Score 24; DB 2; Length 1020;
Best Local Similarity 36.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXXXP 29
Db 453 FAASSAGEYTP 463

RESULT 2
ID 07RAT5 PRELIMINARY; PRT; 214 AA.
AC 07RAT5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transmembrane protein tmp21 precursor.
GN Name=PY06414;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RA PubMed=12368865;
RA Carlton J.M., Angiolini S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kogack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser F.R., Bergman U.W., Vaidya A.B.,
RA van Lin L.H., Jense C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

```

```

CC preliminary data.
DR EMBL; AAB01002169; EAA18629.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008320; F:protein carrier activity; IEA.
DR GO; GO:0006886; P:intracellular protein transport; IEA.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR InterPro; IPR009038; GOLD.
DR Pfam; PF01105; EMP24_GP25L; 1.
DR PROSITE; PSS0866; GOLD; 1.
SQ SEQUENCE 214 AA; 24900 MW; 21A20BFA65BCAF59 CRC64;

Query Match 40.4%; Score 23; DB 2; Length 214;
Best Local Similarity 36.4%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXXXP 29
Db 96 FTFYSGSYSP 106

RESULT 3
ID 09DEB3 PRELIMINARY; PRT; 352 AA.
AC 09DEB3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Inhibin.
OS Name=inh;
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Tada T., Endo M., Hirono I., Takashima F., Aoki T.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the TGF-beta family.
DR EMBL; AB044566; BAB19272.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008083; F:growth factor activity; IEA.
DR GO; GO:0008151; P:cell growth and/or maintenance; IEA.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFp.
DR Pfam; PF00019; TGF_beta; 1.
DR PRINTS; PR00669; INHIBINA.
DR ProDom; PD000357; TGFp; 1.
DR SMART; SM00204; TGFp; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KM Growth factor.
SQ SEQUENCE 352 AA; 39711 MW; ED9CCCEC860F912BD CRC64;

Query Match 40.4%; Score 23; DB 2; Length 352;
Best Local Similarity 36.4%; Pred. No. 9.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXXXP 29
Db 326 FTTSDGYSYP 336

RESULT 4
ID PRTA_PHOAZ STANDARD; PRT; 483 AA.
AC PR2115;
DT 05-JUN-2004 (Ref. 44, Created)
DT 05-JUN-2004 (Ref. 44, Last sequence update)
DT 05-JUN-2004 (Ref. 44, Last annotation update)
DE Secreted alkaline metalloproteinase (EC 3.4.24.-).
GN Name=PRTA;
OS Photobacterium sp. (strain Az29).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```



OC Enterobacteriaceae; Photorhabdus.  
 OX NCBI\_TaxID=229779;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mendes-Cabral C., Montiel R., Simoes N.;  
 RT "Cloning and sequencing of the gene for a metalloprotease with an ABC  
 RT transport system: a protease secreted by Photorhabdus sp. A229  
 RT involved in the inhibition of insect antibacterial peptides.";  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 155-162; 166-197 AND 264-277, AND SUBCELLULAR LOCATION.  
 RA Cherqui A., Mendes-Cabral C., Simoes N.;  
 RT "Two exoproteases from Photorhabdus luminescens.";  
 RL Submitted (SEP-1999) to Swiss-Pro.  
 CC -1- FUNCTION: Involved in the inhibition of insect antibacterial  
 CC peptides.  
 CC -1- COFACTOR: Binds 1 zinc ion and a number of calcium ions per  
 CC subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to peptidase family M10B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
 CC -----  
 DR EMBL: AY511111; AAS19409.1; -  
 DR InterPro: IPR001343; Hemolysin\_Ca\_bind.  
 DR InterPro: IPR006025; Pept\_M\_Zn\_BS.  
 DR PROSITE: PS00330; HEMOLYSIN\_CALCITUM, 1.  
 DR KRM Calcium-binding; Direct protein sequencing; Hydrolyase; Metal-binding;  
 KM Metalloprotease; Repeat; Zinc.  
 FT REPEAT 343 351 GKGXD 1.  
 FT REPEAT 352 360 GKGXD 2.  
 FT REPEAT 361 369 GKGXD 3.  
 FT REPEAT 370 378 GKGXD 4.  
 FT REPEAT 379 387 GKGXD 5.  
 FT METAL 184 184 Zinc (catalytic) (By similarity).  
 FT ACT SITE 185 185 By similarity.  
 FT METAL 188 188 Zinc (catalytic) (By similarity).  
 FT METAL 194 194 Zinc (catalytic) (By similarity).  
 FT METAL 263 263 Calcium 1 (via carbonyl oxygen) (By  
 similarity).  
 FT METAL 266 266 Calcium 1 (By similarity).  
 FT METAL 295 295 Calcium 1 (By similarity).  
 FT METAL 297 297 Calcium 1 (via carbonyl oxygen) (By  
 similarity).  
 FT METAL 298 298 Calcium 2 (via carbonyl oxygen) (By  
 similarity).  
 FT METAL 300 300 Calcium 1 and 2 (By similarity).  
 FT METAL 337 337 Calcium 2 (via carbonyl oxygen) (By  
 similarity).  
 FT METAL 339 339 Calcium 2 (By similarity).  
 SQ SEQUENCE 483 AA; 53482 MW; F402483BD50293A1 CRC64;  
 Query Match 40.4%; Score 23; DB 1; Length 483;  
 Best Local Similarity 36.4%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 19 FXXXXXGXYYF 29  
 Db 135 FDASLTGSYAF 145  
 RESULT 5  
 ID 084F72 PRELIMINARY; PRT; 483 AA.  
 AC 084F72;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Secreted alkaline metalloprotease.  
 GN Name=PRTA;  
 OS Photorhabdus luminescens (Xenorhabdus luminescens).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Photorhabdus.  
 OX NCBI\_TaxID=29488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K122;  
 RA Waterfield N.;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY230749; AAO39136.1; -  
 DR HSRP; O69771; I69K.  
 DR MEROPS; M10.060; -  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0008237; F:metallopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro: IPR001343; Hemolysin\_Ca\_bind.  
 DR InterPro: IPR006026; Peptidase\_N.  
 DR InterPro: IPR006025; Pept\_M\_Zn\_BS.  
 DR Pfam; PF00353; HemolysinCbind; 2.  
 DR PRINTS; PR00313; CABNDGRPT.  
 DR SMART; SM0225; ZINC; 1.  
 DR PROSITE; PS00330; HEMOLYSIN\_CALCITUM, 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN.1.  
 SQ SEQUENCE 483 AA; 53354 MW; 9D9585B915DE09C CRC64;

QY 19 FXXXXXGXYYF 29  
 Db 135 FDASLTGSYAF 145  
 RESULT 6  
 ID AAS19409 PRELIMINARY; PRT; 483 AA.  
 AC AAS19409;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE Secreted alkaline metalloprotease.  
 GN PRTA.  
 OS Photorhabdus sp. A229.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Photorhabdus.  
 OX NCBI\_TaxID=229779;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A229;  
 RA Cabral C.M., Montiel R., Simoes N.;  
 RT "Cloning and sequencing of the gene for a metalloprotease with an ABC  
 RT transport system: a protease secreted by Photorhabdus sp. A229  
 RT involved in the inhibition of insect antibacterial peptides.";  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY511111; AAS19409.1; -  
 KM Metalloprotease; Protease.  
 SQ SEQUENCE 483 AA; 53482 MW; F402483BD50293A1 CRC64;  
 Query Match 40.4%; Score 23; DB 2; Length 483;  
 Best Local Similarity 36.4%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 19 FXXXXXGXYYF 29  
 Db 135 FDASLTGSYAF 145

RESULT 7
Q92LNG PRELIMINARY; PRT; 618 AA.
ID Q92LNG
AC Q92LNG;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein SMC03097.
GN OrderedLocustNames=R02996; ORFNames=SMC03097;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_Taxid=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RC MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe P.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowser L., Bulmerster J., Cadieu E., Capela D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federstepel N.A., Fisher R.F.,
RA Gloux S., Godtje T., Goffeau A., Golding B., Guzy J., Gurjal M.,
RA Hernandez-Lucias I., Hong A., Huizer L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.D., Kallman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA Maury D., Palm C., Peck M.C., Pohl T.M., Portecelle D., Purnelle B.,
RA Rampeger U., Surzycski R., Thebaud P., Vandembol M.,
RA Vorhoeft F.U., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RL "The composite genome of the legume symbiont Sinorhizobium meliloti.";
Science 293:668-672(2001).
[2]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RC MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Guzy J., Bothe G., Ampe P., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godtje T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
RA Pohl T., Portecelle D., Puehler A., Purnelle B., Rampeger U.,
RA Renard C., Thebaud P., Vandembol M., Weidner S., Gallbert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL: AL591792; CAC4575.1; -.
DR InterPro: IPR000184; Bac_surfAg_D15.
DR InterPro: IPR010827; Surf_Ag_VNR.
DR Pfam: PF01103; Bac_surfAc_Ag_1.
DR Pfam: PF07244; Surf_Ag_VNR; 1.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
KV Complete proteome; Hypothetical protein.
SQ SEQUENCE 618 AA; 65655 MW; E51276BB29B1FB78 CRC64;
Query Match 40.4%; Score 23; DB 2; Length 618;
Best Local Similarity 36.4%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 19 FXXXXXGXYPX 29
DB 464 FEASAGGYAF 474
RESULT 8
Q8BPJ1 PRELIMINARY; PRT; 144 AA.
ID Q8BPJ1
AC Q8BPJ1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
library, clone:R130320J01 product:runt related transcription factor 3,
full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

XX NCBI_TaxID=10090;
RN RP
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN RP
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN RP
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RA The RIKEN Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN RP
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
  Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN RP
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagoka S., Sasaki N., Carninci P.,
  Kono H., Akiyama J., Nishi K., Kitanai T., Tashiro H., Itoh M.,
  Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
  Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaishiwagi K.,
  Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,
  Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
  Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
  sequencing pipeline with 384 multiplexed capillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN RP
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RA Adachi J., Aizawa K., Akinura T., Arakawa T., Bono H., Carninci P.,
  Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
  Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
  Hori F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
  Kariha H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
  Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
  Nishi K., Nomura K., Numazaki R., Ono M., Ohnato N., Okazaki Y.,
  Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
  Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
  Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
  Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN RP
RN EMBL; AK053910; BAC3587.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000407; AML1_Runt.
DR PRINTS; PR00967; ONCOGENEAML1.
SQ SEQUENCE 144 AA; 14731 MW; EC4789638ADBD9D7 CRC64;
Query March 38.6%; Score 22; DB 2; Length 144;
Best local similarity 36.4%; Pred. No. 7.3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

Qy 19 FXXXXXGXYP 29  
 Db 66 FYGASGSYGF 76

## RESULT 9

Q6N9E2 PRELIMINARY; PRT; 180 AA.  
 AC Q6N9E2;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Apr-1'-p processing enzyme family protein homolog.  
 GN OrderedLocustNames=RP1607;  
 OS Rhodospseudomonas palustris;  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 CC Bradyrhizobiaceae; Rhodospseudomonas.  
 NCBI\_TaxID=1076;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGA009 / ATCC BAA-98;  
 RX PubMed=14704707; DOI=10.1038/nbt923;  
 RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,  
 RA Land M.L., Pelletier D.A., Bobst C., Torres Y Torres J.L., Peres C.,  
 RA Gibson J.L., Hanson T.E., Bobst C., Torres Y Torres J.L., Peres C.,  
 RA Harrison F.H., Gibson J., Harwood C.S.;  
 RT "Complete genome sequence of the metabolically versatile  
 RT photosynthetic bacterium Rhodospseudomonas palustris.";  
 RL Nat. Biotechnol. 22:55-61(2004).  
 DR EMBL; BX572598; CAB27048.1; -  
 DR InterPro; IPR002589; Alpp.  
 DR Pfam; PF01661; Alpp; 1.  
 DR SMART; SM00506; Alpp; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 180 AA; 18435 MW; ABD79473EDDE394 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 180;  
 Best Local Similarity 36.4%; Pred. No. 9e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYP 29  
 Db 122 FSAISTGVYGF 132

## RESULT 10

CAB27048 PRELIMINARY; PRT; 180 AA.  
 AC CAB27048;  
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DE 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Apr-1'-p processing enzyme family protein homolog.  
 GN RP1607;  
 OS Rhodospseudomonas palustris;  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 CC Bradyrhizobiaceae; Rhodospseudomonas.  
 NCBI\_TaxID=1076;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGA009 / ATCC BAA-98;  
 RX PubMed=14704707;  
 RA Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,  
 RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,  
 RA Gibson J.L., Hanson T.E., Bobst C., Torres Y Torres J.L., Peres C.,  
 RA Harrison F.H., Gibson J., Harwood C.S.;  
 RT "Complete genome sequence of the metabolically versatile  
 RT photosynthetic bacterium Rhodospseudomonas palustris.";  
 RL Nat. Biotechnol. 22:55-61(2004).  
 DR EMBL; BX572598; CAB27048.1; -  
 SQ SEQUENCE 180 AA; 18435 MW; ABD79473EDDE394 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 180;  
 Best Local Similarity 36.4%; Pred. No. 9e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYP 29  
 Db 122 FSAISTGVYGF 132

## RESULT 11

Q8NB08 PRELIMINARY; PRT; 188 AA.  
 AC Q8NB08;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein FLJ90737.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA Iwagai T., Oca T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,  
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,  
 RA Masuho Y., Oho T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,  
 RA Hattori A., Okumura K., Iwayanagi T., Minomiyu K.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK075218; BAC11479.1; -  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0008320; F:protein carrier activity; IEA.  
 DR GO; GO:0006886; P:intracellular protein transport; IEA.  
 DR InterPro; IPR00348; Emp24\_gp25L\_p24.  
 DR InterPro; IPR009038; GOLD.  
 DR Pfam; PF01105; EMP24\_GP25L; 1.  
 DR PROSITE; PS50866; GOLD; 1.  
 SQ SEQUENCE 188 AA; 21233 MW; 15A10CC4BD553DB3 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 188;  
 Best Local Similarity 36.4%; Pred. No. 9.4e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYP 29  
 Db 98 FTASKNGTYKF 108

## RESULT 12

YA09 HUMAN STANDARD; PRT; 215 AA.  
 AC Q9Y3B3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein CGI-109 precursor.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20272150; PubMed=10810093;  
 RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;  
 RT "Identification of novel human genes evolutionarily conserved in  
 RT Caenorhabditis elegans by comparative proteomics.";  
 RL Genome Res. 10:703-713(2000).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic  
 CC reticulum (by similarity). EMP24/GP25L family.  
 CC -1- SIMILARITY: Belongs to the EMP24/GP25L family.  
 CC -1- SIMILARITY: Contains 1 GOLD domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL/AF151867; AAD34104.1; -  
DR InterPro: IPR000348; Emp24\_gp25L\_p24.  
DR InterPro: IPR009038; GOLD.  
DR Pfam: PF01105; EMP24\_GP25L; 1.  
DR PROSITE; PS50866; GOLD; 1.  
KW Endoplasmic reticulum; Hypothetical protein; Signal; Transmembrane.  
FT SIGNAL 1 23 Potential.  
FT CHAIN 24 215 Hypothetical protein CGI-109.  
FT DOMAIN 24 176 Lumenal (Potential).  
FT TRANSMEM 177 199 Potential.  
FT DOMAIN 200 215 Cytoplasmic (Potential).  
FT DOMAIN 38 120 GOLD.  
SQ SEQUENCE 215 AA; 24354 MW; 5C4D5B8E8054857B CRC64;

Query March 38.6%; Score 22; DB 1; Length 215;  
Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Cy 19 FXXXXXGXYP 29  
Db 90 FTSKNGTYKF 100

## RESULT 13

OGNWX3 PRELIMINARY; PRT; 217 AA.

AC 06-JUL-2004 (TREMURel. 27, Created)  
DT 05-JUL-2004 (TREMURel. 27, Last sequence update)  
DE 05-JUL-2004 (TREMURel. 27, Last annotation update)  
DE Hypothetical protein zgc:76996.  
GN Name=zgc:76996;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_Taxid=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshlyuk S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywicki M.I., Skalka U., Smalins D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX Strauberg R.;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL/BC067383; AAH67383.1; -  
DR InterPro: IPR000348; Emp24\_gp25L\_p24.  
DR InterPro: IPR009038; GOLD.  
DR Pfam: PF01105; EMP24\_GP25L; 1.  
DR PROSITE; PS50866; GOLD; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 217 AA; 24555 MW; F885F671CDB86010 CRC64;

Query March 38.6%; Score 22; DB 2; Length 217;  
Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Cy 19 FXXXXXGXYP 29  
Db 91 FSAARNGTYPF 101

## RESULT 14

AAH67383 PRELIMINARY; PRT; 217 AA.

AC AAH67383;  
DT 24-MAY-2004 (TREMURel. 27, Created)  
DT 24-MAY-2004 (TREMURel. 27, Last sequence update)  
DE 24-MAY-2004 (TREMURel. 27, Last annotation update)  
DE Hypothetical protein zgc:76996.  
GN zgc:76996.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_Taxid=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshlyuk S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywicki M.I., Skalka U., Smalins D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX Strauberg R.;

DR EMBL/BC067383; AAH67383.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 217 AA; 24555 MW; F885F671CDB86010 CRC64;

Query March 38.6%; Score 22; DB 2; Length 217;  
Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Cy 19 FXXXXXGXYP 29  
Db 91 FSAARNGTYPF 101

## RESULT 15

06P2V7 PRELIMINARY: PRT: 219 AA.

AC 06P2V7  
 DT 05-JUL-2004 (TRENBLREL. 27, Created)  
 DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)  
 DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)  
 DE Hypothetical protein MGC76316.  
 GN Name=MGC76316;  
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=83364;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska V., Smalins D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Mair M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strauberg R.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, BC064279; AA064279.1; -  
 DR InterPro; IPR000348; EMP24\_GP25L\_P24.  
 DR Pfam; PF01105; EMP24\_GP25L; 1.  
 DR PROSITE; PS50866; GOLD; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 219 AA; 25307 MW; 1D9BDF5626C4C14 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 219;  
 Best Local Similarity 36.4%; Pred.No. 1.1e+03;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 19 FXXXXXGXYX 29  
 DB 93 FTATRNGTYKF 103

Search completed: December 25, 2004, 08:38:30  
 Job time : 60.0095 secs

1115 ruge blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:12:38 ; Search time 13.4399 Seconds  
(without alignments)  
152.967 Million cell updates/sec

Title: US-10-621-787-5

Perfect score: 57  
Sequence: 1 FXXXXXXXXXXXXXXXXXXGXYPX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	22	38.6	266 4	US-09-710-279-2066 Sequence 2066, Ap
2	22	38.6	268 4	US-09-252-991A-18273 Sequence 18273, A
3	22	38.6	292 4	US-09-134-001C-4142 Sequence 4142, Ap
4	22	38.6	305 4	US-09-328-352-6069 Sequence 6069, Ap
5	22	38.6	404 4	US-09-328-352-8182 Sequence 8182, Ap
6	21	36.8	156 4	US-09-158-452A-698 Sequence 698, Ap
7	21	36.8	201 2	US-08-801-740-6 Sequence 6, Appli
8	21	36.8	201 2	US-08-801-740-7 Sequence 7, Appli
9	21	36.8	201 3	US-08-801-740-6 Sequence 6, Appli
10	21	36.8	201 3	US-08-801-740-7 Sequence 7, Appli
11	21	36.8	252 4	US-09-252-991A-30813 Sequence 30813, A
12	21	36.8	450 4	US-09-489-039A-10052 Sequence 10052, A
13	21	36.8	1233 4	US-09-134-000C-4971 Sequence 4971, Ap
14	21	36.8	1416 4	US-09-071-035-404 Sequence 404, App
15	21	36.8	1448 4	US-09-071-035-402 Sequence 402, App
16	20	35.1	63 2	US-08-637-759B-501 Sequence 501, App
17	20	35.1	63 3	US-08-871-355A-501 Sequence 501, App
18	20	35.1	63 3	US-09-201-945-501 Sequence 501, App
19	20	35.1	86 4	US-09-178-093B-34 Sequence 34, Appl
20	20	35.1	110 4	US-09-270-767-44692 Sequence 44692, A
21	20	35.1	142 2	US-08-997-080-47 Sequence 47, Appl
22	20	35.1	142 2	US-08-997-362-47 Sequence 47, Appl
23	20	35.1	142 3	US-08-873-970-47 Sequence 47, Appl
24	20	35.1	142 3	US-09-095-855-47 Sequence 47, Appl
25	20	35.1	142 3	US-08-705-347A-47 Sequence 47, Appl
26	20	35.1	142 3	US-09-324-542-47 Sequence 47, Appl
27	20	35.1	142 4	US-09-205-426-47 Sequence 47, Appl

28	20	35.1	142 4	US-09-200-643-47 Sequence 47, Appl
29	20	35.1	143 3	US-08-685-808-2 Sequence 2, Appli
30	20	35.1	143 3	US-08-505-860C-2 Sequence 2, Appli
31	20	35.1	162 2	US-08-606-143-44 Sequence 44, Appl
32	20	35.1	226 4	US-09-270-767-44619 Sequence 44619, A
33	20	35.1	229 4	US-09-270-767-56939 Sequence 56939, A
34	20	35.1	312 4	US-09-543-681A-5883 Sequence 5883, Ap
35	20	35.1	329 4	US-09-248-796A-16263 Sequence 16263, A
36	20	35.1	336 3	US-08-987-691A-5 Sequence 5, Appli
37	20	35.1	408 4	US-09-252-991A-33131 Sequence 33131, A
38	20	35.1	429 4	US-09-922-501-10 Sequence 4, Appli
39	20	35.1	448 4	US-09-361-443-4 Sequence 3687, Ap
40	20	35.1	462 4	US-09-583-110-3587 Sequence 2, Appli
41	20	35.1	467 4	US-09-361-443-2 Sequence 793, App
42	20	35.1	469 4	US-09-198-452A-793 Sequence 27853, A
43	20	35.1	492 4	US-09-252-991A-27853 Sequence 18, Appl
44	20	35.1	498 3	US-09-232-468A-18 Sequence 24, Appl
45	20	35.1	498 3	US-09-232-468A-24 Sequence 231, App
46	20	35.1	498 3	US-08-686-968C-231 Sequence 52, Appl
47	20	35.1	498 4	US-09-784-984B-52 Sequence 54, Appl
48	20	35.1	498 4	US-09-784-984B-54 Sequence 4, Appli
49	20	35.1	527 4	US-09-103-331-4 Sequence 4, Appli
50	20	35.1	527 4	US-09-631-594-55 Sequence 55, Appl
51	20	35.1	546 2	US-09-067-351-1 Sequence 1, Appli
52	20	35.1	546 3	US-09-360-490-1 Sequence 10, Appl
53	20	35.1	646 3	US-09-336-643A-10 Sequence 16367, A
54	20	35.1	681 4	US-09-248-796A-16367 Sequence 45272, A
55	20	35.1	755 4	US-09-270-767-45272 Sequence 15, Appl
56	20	35.1	1312 4	US-09-489-039A-11544 Sequence 11544, A
57	20	35.1	1430 3	US-09-008-172-2 Sequence 2, Appli
58	20	35.1	1430 3	US-09-210-361-6 Sequence 6, Appli
59	20	35.1	1430 3	US-09-740-274-6 Sequence 6, Appli
60	20	35.1	1430 3	US-08-477-451-4 Sequence 4, Appli
61	20	35.1	3177 2	US-08-525-539A-35 Sequence 35, Appl
62	19	33.3	32 3	US-09-621-976-75072 Sequence 7401, Ap
63	19	33.3	59 4	US-09-621-976-5072 Sequence 5072, Ap
64	19	33.3	62 4	US-09-621-976-5079 Sequence 5079, Ap
65	19	33.3	62 4	US-09-621-976-5080 Sequence 5080, Ap
66	19	33.3	62 4	US-09-621-976-5082 Sequence 5078, Ap
67	19	33.3	63 4	US-09-621-976-75078 Sequence 7507, Ap
68	19	33.3	74 4	US-09-621-976-7511 Sequence 7507, Ap
69	19	33.3	79 4	US-09-621-976-7507 Sequence 7507, Ap
70	19	33.3	79 4	US-09-248-796A-21815 Sequence 21815, A
71	19	33.3	99 4	US-09-621-976-7518 Sequence 7518, Ap
72	19	33.3	99 4	US-09-270-767-33955 Sequence 33955, A
73	19	33.3	101 4	US-09-270-767-54812 Sequence 54812, A
74	19	33.3	101 4	US-09-270-767-58325 Sequence 58325, A
75	19	33.3	104 4	US-09-344-624-7 Sequence 75, Appl
76	19	33.3	107 4	US-08-591-632-85 Sequence 85, Appl
77	19	33.3	108 3	US-08-591-632-85 Sequence 85, Appl
78	19	33.3	108 3	US-09-611-451-75 Sequence 75, Appl
79	19	33.3	108 3	US-09-611-451-75 Sequence 85, Appl
80	19	33.3	108 3	US-09-611-451-85 Sequence 85, Appl
81	19	33.3	110 4	US-09-248-796A-21943 Sequence 21943, A
82	19	33.3	111 4	US-10-162-012-11 Sequence 11, Appl
83	19	33.3	115 4	US-08-956-171E-5217 Sequence 5217, Ap
84	19	33.3	115 4	US-08-781-986A-5217 Sequence 4, Appli
85	19	33.3	122 1	US-08-249-013-4 Sequence 4, Appli
86	19	33.3	122 2	US-08-886-863-4 Sequence 4, Appli
87	19	33.3	122 3	US-09-175-229-4 Sequence 4, Appli
88	19	33.3	122 5	PCT-US95-06764-4 Sequence 680, Ap
89	19	33.3	123 4	US-09-621-976-6980 Sequence 266, Ap
90	19	33.3	126 4	US-09-710-279-2766 Sequence 4973, Ap
91	19	33.3	129 4	US-09-621-976-4973 Sequence 10, Appl
92	19	33.3	130 3	US-09-485-316A-10 Sequence 10, Appl
93	19	33.3	131 4	US-09-107-532A-5741 Sequence 5741, Ap
94	19	33.3	133 4	US-09-583-110-5269 Sequence 5269, Ap
95	19	33.3	145 3	US-09-286-529-21 Sequence 21, Appl
96	19	33.3	153 4	US-09-270-767-46034 Sequence 46034, A
97	19	33.3	154 1	US-07-752-101A-69 Sequence 69, Appl
98	19	33.3	155 3	US-09-134-001C-5566 Sequence 5566, Ap
99	19	33.3	155 4	US-09-270-767-32485 Sequence 32485, A
100	19	33.3	155 4	US-09-270-767-47702 Sequence 47702, A

## ALIGNMENTS

## RESULT 1

US-09-710-279-2066  
; Sequence 2066, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P33480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2066  
; LENGTH: 266  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-2066

## Query Match

Best Local Similarity 38.6%; Score 22; DB 4; Length 266;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 147 FKTTRSGVYTF 157

## RESULT 2

US-09-252-991A-18273  
; Sequence 18273, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18273  
; LENGTH: 268  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18273

## Query Match

Best Local Similarity 38.6%; Score 22; DB 4; Length 268;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 41 FLAVATGTYAF 51

## RESULT 3

US-09-134-001C-4142  
; Sequence 4142, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4142  
; LENGTH: 292  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4142

## Query Match

Best Local Similarity 38.6%; Score 22; DB 3; Length 292;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 173 FKTTRSGVYTF 183

## RESULT 4

US-09-328-352-6069  
; Sequence 6069, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6069  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6069

## Query Match

Best Local Similarity 38.6%; Score 22; DB 4; Length 305;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 52 FTGTAQTYTF 62

## RESULT 5

US-09-328-352-8182  
; Sequence 8182, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 8182  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-8182

## Query Match

Best Local Similarity 38.6%; Score 22; DB 4; Length 404;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;



Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGYXF 29

DB 242 FLKTTAGHYAF 252

## RESULT 6

US-09-198-452A-698

Sequence 698, Application US/09198452A

Patent No. 6559294

GENERAL INFORMATION:

APPLICANT: Griffiths, R.

TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

FILE REFERENCE: 9710-003-999

CURRENT APPLICATION NUMBER: US/09/198,452A

CURRENT FILING DATE: 1998-11-24

NUMBER OF SEQ ID NOS: 6849

SEQ ID NO 698

LENGTH: 156

TYPE: PRT

ORGANISM: Chlamydia pneumoniae

US-09-198-452A-698

Query Match 36.8%; Score 21; DB 4; Length 156;

Best Local Similarity 36.4%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGYXF 29

DB 88 FSLPSSGTYVF 98

## RESULT 7

US-08-801-740-6

Sequence 6, Application US/08801740

Patent No. 5869639

GENERAL INFORMATION:

APPLICANT: Goll, Surya K.

APPLICANT: Bandman, Olga

TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,740

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0189 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 6;

SEQUENCE CHARACTERISTICS:

LENGTH: 201 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1212965

US-08-801-740-6

Query Match 36.8%; Score 21; DB 2; Length 201;

Best Local Similarity 36.4%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGYXF 29

DB 82 FAAMDGTYKF 92

## RESULT 8

US-08-801-740-7

Sequence 7, Application US/08801740

Patent No. 5869639

GENERAL INFORMATION:

APPLICANT: Goll, Surya K.

APPLICANT: Bandman, Olga

TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,740

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0189 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 7;

SEQUENCE CHARACTERISTICS:

LENGTH: 201 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1213221

US-08-801-740-7

Query Match 36.8%; Score 21; DB 2; Length 201;

Best Local Similarity 36.4%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGYXF 29

DB 82 FAAMDGTYKF 92

```
RESULT 9
US-08-801-740-6
; Sequence 6, Application US/08801740
; Patent No. 6130325
; GENERAL INFORMATION:
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,740
; FILING DATE: 14-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0189 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1212965
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-801-740-6
Query Match 36.8%; Score 21; DB 3; Length 201;
Best Local Similarity 36.4%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGYXF 29
Db 82 FAAMDGTYKF 92
RESULT 10
US-08-801-740-7
; Sequence 7, Application US/08801740
; Patent No. 6130325
; GENERAL INFORMATION:
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```

```
ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,740
; FILING DATE: 14-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0189 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1213221
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-801-740-7
Query Match 36.8%; Score 21; DB 3; Length 201;
Best Local Similarity 36.4%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGYXF 29
Db 82 FAAMDGTYKF 92
RESULT 11
US-09-252-991A-30813
; Sequence 30813, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30813
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30813
Query Match 36.8%; Score 21; DB 4; Length 252;
Best Local Similarity 36.4%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGYXF 29
Db 167 FASEDSGVYRF 177
```

RESULT 12  
US-09-489-039A-10052  
Sequence 10052, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 10052  
LENGTH: 450  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10052

Query Match 36.8%; Score 21; DB 4; Length 450;  
Best Local Similarity 36.4%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGYXF 29  
DB 174 PARPDAGVAF 184

RESULT 13  
US-09-134-000C-4971  
Sequence 4971, Application US/09134000C  
Patent No. 6617136  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 4971  
LENGTH: 1233  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-4971

Query Match 36.8%; Score 21; DB 4; Length 1233;  
Best Local Similarity 36.4%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGYXF 29  
DB 990 FTTDEKQYGF 1000

RESULT 14  
US-09-071-035-404  
Sequence 404, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA

ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 404:

Query Match 36.8%; Score 21; DB 4; Length 1416;  
Best Local Similarity 36.4%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGYXF 29  
DB 959 FTTDEKQYGF 969

RESULT 15  
US-09-071-035-402  
Sequence 402, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 402:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1448 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-071-035-402

Query Match 36.8%; Score 21; DB 4; Length 1448;  
Best Local Similarity 36.4%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYP 29  
Db 986 FTTDEKGYSP 996

Search completed: December 25, 2004, 08:42:03  
Job time : 16.4399 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 25, 2004, 08:38:38 / Search time 54.5443 Seconds  
(without alignments)  
204.102 Million cell updates/sec

Title: US-10-621-787-5  
Perfect score: 57  
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXGXXFXFX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1595201 seqs, 359116952 residues

Total number of hits satisfying chosen parameters: 1595201

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

## Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubppaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubppaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubppaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubppaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubppaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubppaa/US10E\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubppaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	38.6	119	US-10-434-156-17	Sequence 17, Appl
2	38.6	136	US-10-434-156-22	Sequence 22, Appl
3	38.6	163	US-09-030-847-2	Sequence 2, Appl1
4	38.6	163	US-10-215-088-2	Sequence 2, Appl1
5	38.6	163	US-10-775-165-2	Sequence 2, Appl1
6	38.6	215	US-09-030-847-4	Sequence 4, Appl1
7	38.6	215	US-10-215-088-4	Sequence 4, Appl1
8	38.6	215	US-10-170-385-289	Sequence 289, Appl
9	38.6	215	US-10-775-165-4	Sequence 4, Appl1
10	38.6	219	US-10-276-774-2150	Sequence 2150, Ap
11	38.6	296	US-10-282-122A-67003	Sequence 67003, A
12	38.6	302	US-10-282-122A-77397	Sequence 77397, A
13	38.6	374	US-10-437-963-159816	Sequence 159816,

14	38.6	636	US-10-425-114-68352	Sequence 68352, A
15	38.6	636	US-10-425-115-274242	Sequence 274242, A
16	38.6	1421	US-10-424-599-275109	Sequence 275109, A
17	38.6	1678	US-10-437-963-126095	Sequence 126095, A
18	38.6	1730	US-10-424-599-275110	Sequence 275110, A
19	38.6	3705	US-10-282-122A-77944	Sequence 77944, A
20	36.8	53	US-10-029-386-29069	Sequence 29069, A
21	36.8	71	US-10-425-115-272794	Sequence 272794, A
22	36.8	71	US-10-425-115-272828	Sequence 272828, A
23	36.8	79	US-10-424-599-240256	Sequence 240256, A
24	36.8	83	US-10-424-599-168732	Sequence 168732, A
25	36.8	100	US-10-425-115-364312	Sequence 364312, A
26	36.8	128	US-10-466-242-51	Sequence 51, Appl1
27	36.8	131	US-10-767-701-42874	Sequence 42874, A
28	36.8	150	US-10-437-963-154089	Sequence 154089, A
29	36.8	156	US-10-289-762-698	Sequence 698, App
30	36.8	156	US-10-282-122A-54708	Sequence 54708, A
31	36.8	156	US-10-425-115-333955	Sequence 333955, A
32	36.8	158	US-10-425-115-333946	Sequence 333946, A
33	36.8	165	US-10-425-114-44317	Sequence 44317, A
34	36.8	201	US-09-833-245-2073	Sequence 2073, Ap
35	36.8	201	US-09-833-245-2075	Sequence 2075, Ap
36	36.8	201	US-09-833-245-2076	Sequence 2076, Ap
37	36.8	201	US-10-039-865-6	Sequence 6, Appl1
38	36.8	201	US-10-039-865-7	Sequence 7, Appl1
39	36.8	201	US-10-060-255-59	Sequence 59, Appl
40	36.8	201	US-10-060-255-68	Sequence 68, Appl
41	36.8	201	US-10-060-255-69	Sequence 69, Appl
42	36.8	201	US-10-408-765A-717	Sequence 717, Appl
43	36.8	201	US-10-881-088-59	Sequence 59, Appl
44	36.8	201	US-10-881-088-68	Sequence 68, Appl
45	36.8	201	US-10-881-088-69	Sequence 69, Appl
46	36.8	208	US-10-425-115-321090	Sequence 321090, A
47	36.8	210	US-10-335-977-5883	Sequence 5883, Ap
48	36.8	215	US-10-335-977-5884	Sequence 5884, Ap
49	36.8	215	US-10-335-977-5885	Sequence 5885, Ap
50	36.8	233	US-10-424-599-165427	Sequence 165427, A
51	36.8	238	US-10-282-122A-58562	Sequence 58562, A
52	36.8	277	US-10-425-114-43315	Sequence 43315, A
53	36.8	277	US-10-425-114-43316	Sequence 43316, A
54	36.8	300	US-09-769-736-86	Sequence 736, Ap
55	36.8	312	US-10-425-115-293160	Sequence 293160, A
56	36.8	312	US-10-425-115-293161	Sequence 293161, A
57	36.8	336	US-10-425-115-190446	Sequence 190446, A
58	36.8	356	US-10-425-115-293155	Sequence 293155, A
59	36.8	356	US-10-425-115-293156	Sequence 293156, A
60	36.8	376	US-10-425-114-55341	Sequence 55341, A
61	36.8	397	US-10-437-963-132936	Sequence 132936, A
62	36.8	408	US-10-781-014-462	Sequence 462, App
63	36.8	412	US-10-425-114-43315	Sequence 43315, A
64	36.8	420	US-10-369-493-11376	Sequence 11376, A
65	36.8	420	US-10-369-493-11377	Sequence 11377, A
66	36.8	420	US-10-369-493-11378	Sequence 11378, A
67	36.8	420	US-10-369-493-11379	Sequence 11379, A
68	36.8	421	US-10-369-493-11380	Sequence 11380, A
69	36.8	421	US-10-369-493-11381	Sequence 11381, A
70	36.8	422	US-10-369-493-11382	Sequence 11382, A
71	36.8	429	US-10-369-493-11383	Sequence 11383, A
72	36.8	441	US-10-369-493-11384	Sequence 11384, A
73	36.8	466	US-10-424-599-260503	Sequence 260503, A
74	36.8	473	US-10-425-1107-69	Sequence 69, Appl
75	36.8	501	US-10-425-1107-69	Sequence 69, Appl
76	36.8	508	US-10-342-224-42	Sequence 42, Appl
77	36.8	525	US-10-424-599-192279	Sequence 192279, A
78	36.8	600	US-10-425-115-354835	Sequence 354835, A
79	36.8	661	US-10-369-493-10446	Sequence 10446, A
80	36.8	700	US-10-437-963-162308	Sequence 162308, A
81	36.8	803	US-10-437-963-154762	Sequence 154762, A
82	36.8	811	US-10-437-963-123314	Sequence 123314, A
83	36.8	845	US-10-437-963-163017	Sequence 163017, A
84	36.8	923	US-10-437-963-163015	Sequence 163015, A
85	36.8	1072	US-10-437-963-180020	Sequence 180020, A
86	36.8	1119	US-10-437-963-114130	Sequence 114130, A

```
87      21      36.8      1393      16      US-10-437-963-114133      Sequence 114133,
88      21      36.8      1416      9      US-09-071-035-404      Sequence 404, App
89      21      36.8      1416      14      US-10-206-576-404      Sequence 404, App
90      21      36.8      1448      9      US-09-071-035-402      Sequence 402, App
91      21      36.8      1448      14      US-10-206-576-402      Sequence 402, App
92      21      36.8      1732      16      US-10-437-963-135872      Sequence 135872,
93      20      35.1      20      14      US-10-123-101-84      Sequence 84, App1
94      20      35.1      20      14      US-10-310-734-114      Sequence 114, App1
95      20      35.1      55      14      US-10-029-386-28388      Sequence 28388, A
96      20      35.1      57      15      US-10-424-599-205858      Sequence 205858,
97      20      35.1      59      17      US-10-425-115-309502      Sequence 309502,
98      20      35.1      63      15      US-10-424-599-220167      Sequence 220167,
99      20      35.1      66      15      US-10-424-599-217628      Sequence 217628,
100     20      35.1      71      16      US-10-437-963-155241      Sequence 155241,
```

## ALIGNMENTS

```
RESULT 1
US-10-434-156-17
; Sequence 17, Application US/10434156
; Publication No. US20040019006A1
; GENERAL INFORMATION:
; APPLICANT: TORAY Industries Inc.
; APPLICANT: Yoshihide HAYASHIZAKI
; APPLICANT: Mamoru KAMITVA
; APPLICANT: Tomohiko SUZUKI
; APPLICANT: Hiroe HIROKAWA
; APPLICANT: Kaoru NAKAO
; APPLICANT: Toshiaki TANAKA
; TITLE OF INVENTION: No. US20040019006A1el genes relating to pain and use of the genes
; FILE REFERENCE: 1254-0222P
; CURRENT APPLICATION NUMBER: US/10/434,156
; PRIOR FILING DATE: 2003-05-09
; PRIOR FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 2.1
; SEQ ID NO 17
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-434-156-17

Query Match      38.6%; Score 22; DB 15; Length 119;
Best Local Similarity 36.4%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      19      FXXXXXGXYXF 29
DB      10      FTASRNGTYKF 20

RESULT 2
US-10-434-156-22
; Sequence 22, Application US/10434156
; Publication No. US20040019006A1
; GENERAL INFORMATION:
; APPLICANT: TORAY Industries Inc.
; APPLICANT: Yoshihide HAYASHIZAKI
; APPLICANT: Mamoru KAMITVA
; APPLICANT: Tomohiko SUZUKI
; APPLICANT: Hiroe HIROKAWA
; APPLICANT: Kaoru NAKAO
; APPLICANT: Toshiaki TANAKA
; TITLE OF INVENTION: No. US20040019006A1el genes relating to pain and use of the genes
; FILE REFERENCE: 1254-0222P
; CURRENT APPLICATION NUMBER: US/10/434,156
; PRIOR FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/378,955
; PRIOR FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 28
```

```
; SOFTWARE: PatentIn version 2.1
; SEQ ID NO 22
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-434-156-22
```

```
Query Match      38.6%; Score 22; DB 15; Length 136;
Best Local Similarity 36.4%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      19      FXXXXXGXYXF 29
DB      10      FTASKNGTYKF 20
```

```
RESULT 3
US-09-030-847-2
; Sequence 2, Application US/09030847A
; Patent No. US20020111472A1
; GENERAL INFORMATION:
; APPLICANT: NI, Jian
; TITLE OF INVENTION: T1-R Ligand III
; FILE REFERENCE: PF357
; CURRENT APPLICATION NUMBER: US/09/030,847A
; CURRENT FILING DATE: 1998-02-26
; EARLIER APPLICATION NUMBER: 60/039,483
; EARLIER FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-030-847-2
```

```
Query Match      38.6%; Score 22; DB 9; Length 163;
Best Local Similarity 36.4%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      19      FXXXXXGXYXF 29
DB      89      FTASKNGTYKF 99
```

```
RESULT 4
US-10-215-088-2
; Sequence 2, Application US/10215088
; Publication No. US20030069409A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: T1-R Ligand III
; FILE REFERENCE: PF357C1
; CURRENT APPLICATION NUMBER: US/10/215,088
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/030,847
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/039,483
; PRIOR FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-215-088-2
```

```
Query Match      38.6%; Score 22; DB 14; Length 163;
Best Local Similarity 36.4%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      19      FXXXXXGXYXF 29
```

Db 89 FTASKNGTYKF 99

RESULT 5

US-10-775-165-2  
Sequence 2, Application US/10775165  
Publication No. US20040138443A1  
GENERAL INFORMATION:  
APPLICANT: NI et al.  
TITLE OF INVENTION: TI-R Ligand III  
FILE REFERENCE: PF357C1  
CURRENT APPLICATION NUMBER: US/10/775,165  
CURRENT FILING DATE: 2004-02-11  
PRIOR APPLICATION NUMBER: US/10/215,088  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 09/030,847  
PRIOR FILING DATE: 1998-02-26  
PRIOR APPLICATION NUMBER: 60/039,483  
PRIOR FILING DATE: 1997-02-28  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Version 3.1  
SEQ ID NO 2  
LENGTH: 163  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-775-165-2

Query Match 38.6%; Score 22; DB 16; Length 163;  
Best Local Similarity 36.4%; Pred. No. 6.8e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYKF 29  
Db 89 FTASKNGTYKF 99

RESULT 6

US-09-030-847-4  
Sequence 4, Application US/09030847A  
Patent No. US2002011472A1  
GENERAL INFORMATION:  
APPLICANT: NI, Jian  
TITLE OF INVENTION: TI-R Ligand III  
FILE REFERENCE: PF357  
CURRENT APPLICATION NUMBER: US/09/030,847A  
CURRENT FILING DATE: 1998-02-26  
EARLIER APPLICATION NUMBER: 60/039,483  
EARLIER FILING DATE: 1997-02-28  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 215  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-030-847-4

Query Match 38.6%; Score 22; DB 9; Length 215;  
Best Local Similarity 36.4%; Pred. No. 8.8e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYKF 29  
Db 89 FTASKNGTYKF 99

RESULT 7

US-10-215-088-4  
Sequence 4, Application US/10215088  
Publication No. US20030069409A1  
GENERAL INFORMATION:  
APPLICANT: NI et al.  
TITLE OF INVENTION: TI-R Ligand III  
FILE REFERENCE: PF357C1

CURRENT APPLICATION NUMBER: US/10/215,088

CURRENT FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 09/030,847  
PRIOR FILING DATE: 1998-02-26  
PRIOR APPLICATION NUMBER: 60/039,483  
PRIOR FILING DATE: 1997-02-28  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Version 3.1  
SEQ ID NO 4  
LENGTH: 215  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-215-088-4

Query Match 38.6%; Score 22; DB 14; Length 215;  
Best Local Similarity 36.4%; Pred. No. 8.8e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYKF 29  
Db 89 FTASKNGTYKF 99

RESULT 8

US-10-170-385-289  
Sequence 289, Application US/10170385  
Publication No. US20030203372A1  
GENERAL INFORMATION:  
APPLICANT: Ward, Neil Raymond  
APPLICANT: Mundy, Christopher Robert  
APPLICANT: Kan, On  
APPLICANT: Harris, Robert Alan  
APPLICANT: White, Jonathan  
APPLICANT: Binley, Katie Mary  
APPLICANT: Rayner, William Nigel  
APPLICANT: Naylor, Stuart  
APPLICANT: Kingsman, Susan Mary  
APPLICANT: Krige, David  
TITLE OF INVENTION: ANALYSIS METHOD  
FILE REFERENCE: 53268200100  
CURRENT APPLICATION NUMBER: US/10/170,385  
CURRENT FILING DATE: 2002-06-12  
PRIOR APPLICATION NUMBER: PCT/GB02/01662  
PRIOR FILING DATE: 2002-04-08  
PRIOR APPLICATION NUMBER: PCT/GB01/05458  
PRIOR FILING DATE: 2001-12-10  
NUMBER OF SEQ ID NOS: 549  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 289  
LENGTH: 215  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-170-385-289

Query Match 38.6%; Score 22; DB 14; Length 215;  
Best Local Similarity 36.4%; Pred. No. 8.8e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYKF 29  
Db 90 FTASKNGTYKF 100

RESULT 9

US-10-775-165-4  
Sequence 4, Application US/10775165  
Publication No. US20040138443A1  
GENERAL INFORMATION:  
APPLICANT: NI et al.  
TITLE OF INVENTION: TI-R Ligand III  
FILE REFERENCE: PF357C1  
CURRENT APPLICATION NUMBER: US/10/775,165  
CURRENT FILING DATE: 2004-02-11

```

; PRIOR APPLICATION NUMBER: US/10/215,088
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/030,847
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/039,483
; PRIOR FILING DATE: 1997-02-28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-775-165-4

Query Match      38.6%; Score 22; DB 16; Length 215;
Best Local Similarity 36.4%; Pred. No. 8.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXYXF 29
Db      89 FTASKNGTYKF 99

RESULT 10
; Sequence 2150, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2150
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-276-774-2150

Query Match      38.6%; Score 22; DB 15; Length 219;
Best Local Similarity 36.4%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXYXF 29
Db      93 FTASKNGTYKF 103

RESULT 11
; Sequence 67003, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67003
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Pasteurella multocida
; US-10-282-122A-67003

Query Match      38.6%; Score 22; DB 15; Length 296;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXYXF 29
Db      276 FTAKVGSYKF 286

RESULT 12
; Sequence 77397, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```



```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27/257,931
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77397
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77397

Query Match          38.6%; Score 22; DB 15; Length 302;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXGXYP 29
DB      277 FTKAVTGAYPF 287

RESULT 13
US-10-437-963-159816
; Sequence 159816, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159816
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59155C.1.pep
US-10-437-963-159816

Query Match          38.6%; Score 22; DB 16; Length 374;
Best Local Similarity 36.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXGXYP 29
DB      207 FSAHGDTYSP 217

RESULT 14
US-10-425-114-68352
; Sequence 68352, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingtong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B

APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53131)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68352
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZWFLMO17067D05_FLI.pep
US-10-425-114-68352

Query Match          38.6%; Score 22; DB 15; Length 636;
Best Local Similarity 36.4%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXGXYP 29
DB      500 FVASSAGLYRF 510

RESULT 15
US-10-425-115-274242
; Sequence 274242, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 274242
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_181699C.1.pep
US-10-425-115-274242

Query Match          38.6%; Score 22; DB 17; Length 636;
Best Local Similarity 36.4%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXGXYP 29
DB      500 FVASSAGLYRF 510

Search completed: December 25, 2004, 09:22:22
Job time : 59.5443 secs
```

**This Page Blank (uspto)**